Visual Interpretation of Lambertian Surface Deformation

Richard Michael Cameron-Jones

Ph.D.

University of Edinburgh

1990

Abstract

The major topic of this thesis is the interpretation (as a three-dimensional velocity field) of the changing intensity pattern induced by a smoothly deforming Lambertian surface of uniform albedo illuminated by a distant point light source. A constraint is derived which shows how the changing intensity pattern induced by such a deforming surface is locally constrained by the three-dimensional motion of that surface. This constraint, the "Intensity Rate Constraint", a partial differential equation in the normal component of surface velocity, contains no terms relating to the tangential components of surface velocity, hence the problem of determining the three-dimensional motion is *ill-posed.* The application of an additional constraint on the motion, (implemented in the form of a stretch-based regulariser) is proposed. This enables certain psychologically significant classes of three-dimensional velocity field over the surface to be estimated veridically from the image intensity rate, the velocity field along the boundary and static information. This technique is successfully tested on synthetic data in experiments requiring at least ten times greater accuracy in intensity measurement than is commonly available. The thesis concludes with a suggested technique for the interpretation of smoothly deforming space-curve motion.

Acknowledgements

I would like to thank the following for their help in the preparation of this thesis: Drs. R.B.Fisher and J.C.T.Hallam for supervision, H.W.Hughes for proof-reading, D.Howie for photography, S.J.S.Cranefield for typographic assistance, the SERC for funding, and Prof. J.A.M.Howe for extending the use of the department's facilities when I was no longer funded.

I would also like to generally thank the following past and present members of the department for putting up with me as co-officee, squash opponent or dining companion: J.M.Lewis, A.K.Robertson, L.D.Cai, G.A.Roberts, C.P.Williams, R.J.Caley, M.J.Jones (with whom I was often mistaken!), M.S.Wilson, D.Bental and J.C.Salmon.

Penultimately I would like to mention those former members of the Edinburgh University Chess Club to whom I have lost several games, not just of chess: M.P.Stewart, A.J.Coull, J.I.M.Grant and D.Ward.

Finally, thanks to my parents.

Declaration

I declare that this thesis has been composed by myself and that the work described in it is my own.

R. M. Cameron

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Starting with a Blank Sheet of Paper

The major theme of this work is the interpretation, as a three-dimensional velocity field, of the changing intensity pattern induced by a deforming Lambertian surface. The broader topic of general non-rigid surface motion was narrowed to that addressed by seeking the simplest example of such motion which previous techniques were unable to interpret correctly as a three-dimensional velocity field. Bending a blank sheet of paper seemed an apparently simple example in that the motion of the corners could theoretically be found by a stereo-motion technique, but estimation of the motion of the sheet edge, being that of a non-rigid three-dimensional curve, does not fit into a problem category previously considered, nor does the estimation of the motion over the surface of the sheet. Thus, in contrast to the usual metaphor in design, the "blank sheet of paper" was a specific rather than a general starting point, leading to the broader topic of interest - fully visible, smoothly deforming and shaded surface patches bounded by edges with identifiable feature points, hereafter termed "corners" (as in (Brady, 1987)).

The problem of estimating the motion over the smoothly shaded deforming surface seemed of greater interest than that of estimating the motion along the deforming edges, and was itself made more specific by considering the case in which the surface is presumed to be Lambertian, with known albedo, illuminated by a known distant point light source. (These are common assumptions

in *shape from shading*, which is the standard computer vision approach to determining the static shape of a smoothly shaded surface, such as the blank sheet of paper considered, from an input intensity image.) As the corners' three-dimensional motion is presumed to be known from their two-dimensional motion in a stereo pair, the static shape of the edges may also be assumed to be known from the stereo information, and hence there is sufficient information for the instantaneous surface shape to be determined by application of a shape from shading technique within a bounding contour, as analysed in (Blake et al, 1985).

However, this combination of surface reflectivity and illumination has recently been used (Verri & Poggio, 1987) as an example where, even for a rigid rotation, the standard methods of motion computation would fail (except in the case of rotation around an axis parallel to the illumination direction) because the standard presumed constraint is broken. Whereas the breakdown of this constraint was then used in (Verri & Poggio, 1987) to support the argument for a "qualitative" approach to interpreting images induced by moving surfaces (as developed further in (Verri et al, 1989)), this work explores the alternative of evaluating what constraint does hold and how it could be used. While the mathematical notation introduced in the next chapter is required to understand the detail of the results, a brief verbal description with an example will be given here.

The derived "Intensity Rate Constraint," which relates the change in intensity at an image point to the motion of the surface, contains terms relating to the normal velocity component at the point, but not the tangential components of the velocity. Thus, for the example patch of toroidal surface illustrated in Figure 1–1 (viewed along the torus's axis of rotational symmetry), expanding, rotating and translating, so as to have the instantaneous velocity field depicted in Figure 1–2, the tangential components of velocity at a point within the region are unobservable from the local intensity change. However, it is possible to estimate the full three-dimensional velocity field by the application of regularisation, for certain expected types of motion field. The use of stretch-based

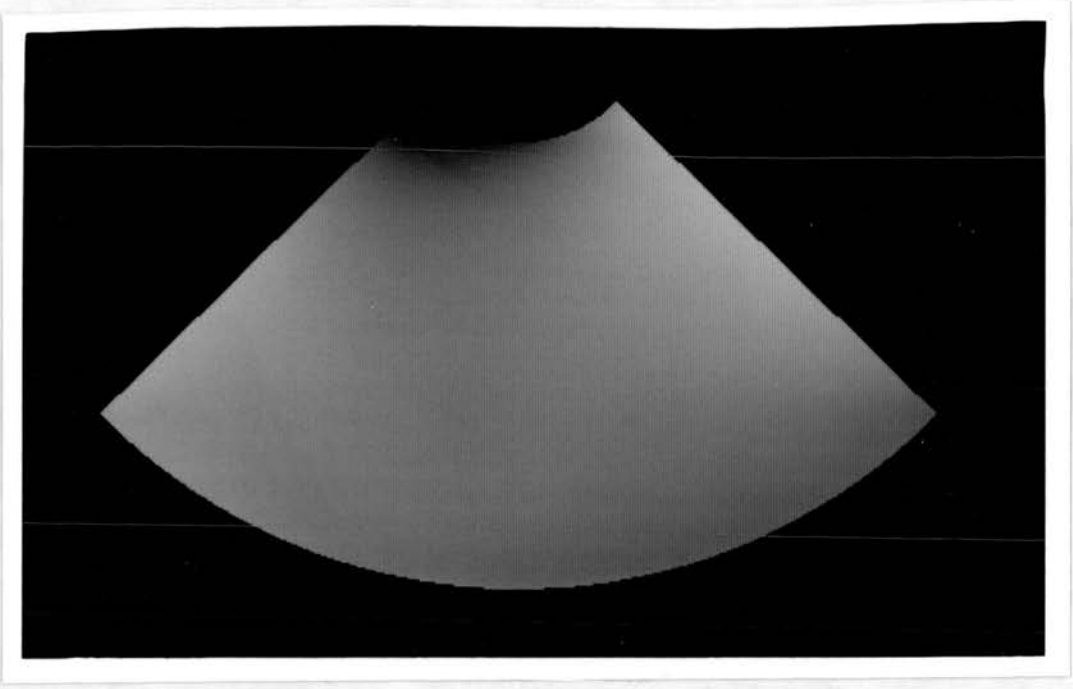

Figure 1–1: Toroidal region intensity image

regularisers is proposed for their ability to estimate rigid or bending motions, and in a limiting case pure stretching motions such as that depicted in Figure 1–2. The method is successfully tested on synthetic data representing first a reduced dimension case in which the problem can be reduced to that of estimating planar curve motion and then the full case including the example torus motion.

It should immediately be made clear that any change in image intensity due to the rotation of a uniform albedo Lambertian surface must be generally expected to be much smaller in magnitude than the change in image intensity which would be caused by the same motion of the surface if it were strongly visually textured. In short, the effect being considered here is rather more subtle than that which is normally considered in methods for interpreting changing intensity images, and consequently practical testing of the technique suggested in this thesis would require more accurate measurement of the intensity rate than is currently available with standard technology. Reproduction of the experiments synthesised would for example require an intensity accuracy of better than 0.1, in images with a maximum intensity of 255, to demonstrate the mag-

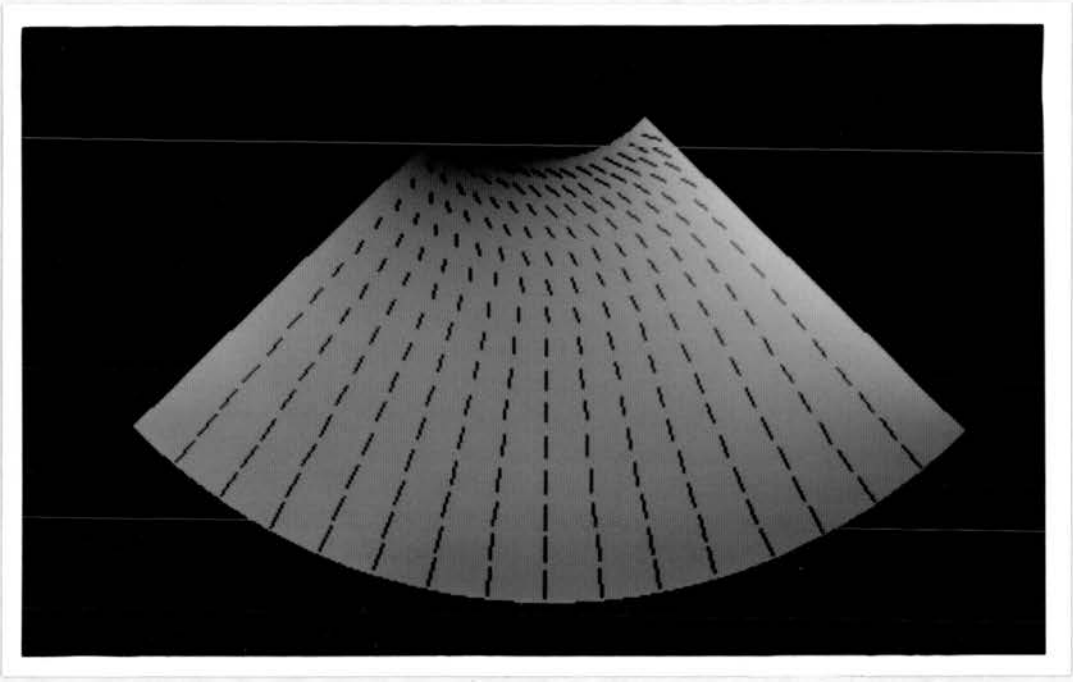

Figure 1–2: Toroidal region expansion, rotation and translation - velocity field superposed on intensity image

nitude of the velocity error due to the spatial discretisation in the solution – about 1% in the main instances tested.

The remaining sections of this chapter describe the motivation behind the choice of general research topic while placing it in the context of a hypothetical complete vision system, the "computational" approach due to David Marr (Marr, 1982), which this work attempts to follow, and the outline of the work in the rest of the thesis.

1.2 Motivation

The idea of investigating methods for determining the three-dimensional velocity of non-rigid surfaces arose from the intersection of the two separate areas of interest - the vision of non-rigid bodies and the vision of dynamic scenes.

1.2.1 Vision of Non-Rigid Bodies

In general, computer vision research has tended to concentrate on the domain of rigid bodies. Whilst previous systems have modelled (Brooks, 1981) and recognised (Fisher, 1989) flexibly attached rigid bodies, and systems which have analysed natural scenes, *e.g.* VISIONS (Hanson & Riseman, 1978) have included non-rigid bodies, *e.g.* trees, these systems have not explicitly modelled the non-rigidity of bodies.

Thus, even if considerable progress is made on the vision problems which most researchers are addressing, the resulting vision system(s) may only be truly capable in a world consisting exclusively of rigid bodies. Yet, the world contains many objects for which the non-rigidity is significant, *e.g.* most clothing. Consequently, the system might not even be truly capable in some industrial situations involving the manufacture of such non-rigid objects, *e.g.* garment manufacture. It should be noted that the domain of non-rigid objects seems more difficult even for human visual systems in terms of recognition tasks. Thus, for example, a heap of previously unseen rigid objects such as hand tools are generally easier to recognise than a heap of previously unseen non-rigid objects such as clothes.

Vision of non-rigid bodies may enable some high level information of interest to be derived - information about properties of materials and supports as in (Pentland & Williams, 1989), where an object's three-dimensional motion and instantaneous shape over a period of time are used to derive material properties such as stiffness (up to a scale constant relating to the object's den-

sity) assuming homogeneity of the material involved. An extreme example of possible inference could be where a sheet is used as a dust cover over an item of furniture. A human observer could infer certain aspects of the shape of the item of furniture, which is acting as a support to the dust cover (enabling a rather weak recognition of a completely occluded object assuming a sufficiently limited domain of furniture!) and might infer some measure of stiffness in relation to density for the cover from the way in which it hangs over the edge of the support - using in both inferences suppositions about gravity. The remainder of this section is concerned with the following major issues in non-rigid body vision - the types of (visually appreciable) non-rigidity, and the visual observation of non-rigidity.

1.2.1.1 Types of Visually Appreciable Non-Rigidity

In the context of a computer vision system, the only non-rigidity of interest is that which is visually appreciable. Thus whilst a materials scientist could correctly observe that all materials deform under stress, and hence rigid objects do not exist, it is generally acceptable to consider the concept of rigidity versus non-rigidity. The (imprecise) distinction essentially arises from the viewing scale (*e.g.* not using a microscope to observe strain) and the level of stresses expected in the environment under consideration.

Further, it is generally accepted that non-rigid objects may be categorised; thus for example in (Pickett, 1984) when considering the modelling of robots and their environment, the following three categories are defined:

1. Elastic objects : "those which are capable of recovering their size and shape after being deformed" *e.g.* hoses, cables and wires on a robot.

2. Plastic objects : "those which are capable of being deformed continuously and permanently" *e.g.* sheet metal to be shaped by stamping in a die.

3. Mechanisms : "those which are made up of essentially rigid components linked together by joints which permit the components to change their

positions with respect to one another" *e.g.* robots (excluding the control hoses, cables and wires as in 1 above).

For the purposes of this work, "mechanisms" are not of interest, being a category which is already competently handled by scene understanding systems (*e.g.* (Fisher, 1989)). Further, it will generally be presumed that the motion of the body under consideration is visibly influenced by a tendency to retain properties of its shape, hence the "elastic" category is more significant than the "plastic".

The broad category of what one might consider "significantly non-rigid objects" may be refined further, as a psychological study of human observation of objects in motion (Jansson & Johansson, 1973) suggests that a distinction should be made between *bending* and *stretching* modes of non-rigidity, corresponding to deformations in which internal lengths (*e.g.* those along a surface) are preserved or altered respectively. The study examined the experimental subjects' reported perception of two-dimensional stimuli in which an outline quadrangle (initially square) was observed to change in side length and direction. The subjects were given three major categories of three-dimensional motion (based on the results of previous work suggesting their significance) - rotation, bending and stretching - and were requested to thus (or otherwise) categorise the perceived stimuli. The results showed that the subjects preferred the rigid rotation explanation, if compatible with the stimulus, over the bending explanation, which was in turn preferred to the stretching explanation, if compatible with the stimulus. This accords with a general principle of "minimum object change", which is consistent with the preference for explaining two-dimensional stimuli which may be perceived as stretching in the plane, as rigid rotation in three dimensions.

This preference for explaining a two-dimensional quadrangle as a three-dimensional object has also been considered in the static case where there is a preference for "rectangularity" - "a single quadrilateral tends always to be perceived as a rectangle" (Gang & Saburo, 1989). It would seem that this principle is broken in the work on the changing quadrangle, perhaps an

indication of the strength of the principle of minimum object change, and the significance of the description categories. In the rest of this chapter, bodies which can sustain reversible bending, but not stretching, will be referred to as "flexible", and those which can sustain reversible stretching will be referred to as "elastic". This distinction is at the "naive physics" level and not the informed scientific level at which bending implies stretching.

1.2.1.2 Visual Observation of Non-Rigidity

The observation of an object in non-rigid motion, *e.g.* a sheet of paper bending, a balloon expanding, a flag or a sail flapping in the wind, is the direct and obvious method of perceiving non-rigidity; however as mentioned earlier, it is also possible that the static shape of an object somehow suggests that it is non-rigid. Thus, digressing from the main theme of non-rigid motion for a final few comments on the static case seems appropriate.

It should be noted that any static observations regarding non-rigidity are only suggestive, thus for example frozen washing on the line may be the same shape as before freezing when it was not rigid. Hence if the shape suggests non-rigidity, frozen washing will appear non-rigid, as indeed it does to a casual human observer (though this could be due to knowledge regarding washing rather than the observation of the shape). The possibility of assessing rigidity suggestiveness from shape could present an interesting criterion in such areas as judging the accuracy of statues of clothed people! If static observation of drapery is considered further, it seems that folds are critical to any static suggestiveness and contain information about the internal forces and hence external loading and support. Perhaps the simplest example of this would be to consider a sheet draped over a vertical pole - the sheet will hang from the pole with the folds running up to the point of support.

If one considers flexible and elastic bodies of different levels of dimensionality it seems that they may have characteristic aspects of shape in certain situations of support and loading. For example, a "1-D" flexible body with

low stiffness, such as a piece of string, will adopt a catenary shape under uniform loading, when held by two point supports at the ends. When the string has intermediate supports, their positions will be evident from the changes in curvature. At the opposite end of the categories, a "3-D" elastic body, such as a cube of rubber used to isolate a mass from minor vibrations, will bulge characteristically under a compression load.

Returning to the matter of non-rigid motion interpretation, it should be mentioned that relevant work in the field of computer vision will be reviewed in the Background chapter, but one of its major attributes should be summarised here as an explanation of the choice made in this thesis to estimate the motion of a *smoothly* shaded surface. Previous work, usually an extension of work on rigid motion interpretation, has generally assumed that the instantaneous two-dimensional velocities (the "optic flow"), or correspondences (in the case of discrete motion), which are the projection of the three-dimensional motion are given as inputs; thus seem to be either presuming that the surface is heavily visually textured, or overlooking the major theoretical problem in the smoothly shaded case.

1.2.2 Vision of Dynamic Scenes

The most generally influential work in the area of vision research is that of David Marr. His book (Marr, 1982) proposed the need for a "computational approach" to the study of vision, as will be described in the next section, and suggested a representational framework for what was considered the "quintessential fact of computer vision - that it tells about shape and space and spatial arrangement". In this framework, the visual system processes the input intensity image through three successive levels of representation, the Primal sketch representing image properties, the $2\frac{1}{2}$D sketch representing surface orientations and depths in a viewer-centred coordinate frame, and the 3D model representation using an object-centred frame.

It is evident that this framework, aimed at representing surface shape as derived from the input, does not make explicit facts about the motion of visible surfaces, and indeed Marr comments that facts other than shape, such as motion, "could be hung off a theory in which the main job of vision was to derive a representation of shape". If the sole purpose of vision is taken to be the recognition of objects, and it is believed that object recognition is performed by recognising shapes and spatial layout, then clearly Marr's framework is sufficient, and in practice many complete vision systems, such as (Fisher, 1989), have recognition by this means as their unique purpose. Further, Marr's framework may also be used for other purposes, such as planning actions which do not require knowledge of how the spatial layout is changing.

However, in a dynamic world, where objects are in relative motion, it seems inadequate to rely upon a static representation despite its considerable utility for recognition in the classic sense, one of the most significant tasks of vision. An example which illustrates a point of ambiguity in the sense in which recognition is considered is that of a cricket batsman "recognising" the "type" of ball bowled, *e.g.* medium pace. This sense of recognition relates to the motion of the ball, as opposed to the classic sense which would involve perhaps identifying that it was the modern low seam ball, which might also in practice be easier to determine from the trajectory of the ball than any instantaneous view, other than the extremely close one the batsman prefers to avoid!

The most straightforward augmentation of Marr's scheme to include motion at the middle level seems to be the addition to the $2\frac{1}{2}$D sketch of estimates of the three-dimensional velocity at points on the surface. It should perhaps be noted that Marr did not ignore the issue of motion, but considered it as it related to the estimation of surface shape - "structure from motion"; thus motion was used in the construction of the $2\frac{1}{2}$D sketch, but not encoded within it. The work done in this thesis addresses a specific problem within the issue of estimating the velocities for this augmented or "dynamic $2\frac{1}{2}$D sketch". Similarly to the static information in the sketch being derived by a variety of modules depending upon the surface characteristics, (*e.g.* deriving shape from shading and shape from

texture although mathematically related requires different surface reflectivity), the motion information may be derived by a variety of modules depending upon the surface characteristics. If the analysis of the motion of heavily textured surfaces viewed stereoscopically, stereo-motion, is considered as the dynamic analogue of stereo, the topic of research in this thesis may be regarded (if the eminent workers in the field will forgive the presumption) as a dynamic analogue of "shape from shading".

The augmentation of Marr's scheme at the higher level is not considered in the rest of this thesis, but it would appear from the psychological work previously considered (Jansson & Johansson, 1973) that the 3-D model representation should include information as to whether the body was static, in rigid motion (specifying translational and rotational velocities), undergoing a bending motion or undergoing a stretching motion. The form of high-level description appropriate for the latter cases, apart from such simple cases as uniform expansion, is unclear.

1.3 The Computational Approach

Marr, in *e.g.* (Marr, 1982), took the stance that the human vision system should be analysed as an information processing system and that any component process of a vision system should be analysed at three separate levels. The first is the "computational theory", in which the outputs of the process are identified and the constraints which govern the outputs and relate them to the inputs (thus determining the process) are formulated. The second level is that of the representation of inputs and outputs and the algorithm used to determine the outputs from the inputs. The third level is that of the implementation used.

In this thesis the major emphasis is placed upon determining the constraints at the computational level, as the aspects more related to implementation on a digital computer are comparatively routine. It should be noted that although generally motivated by the ability of humans to interpret non-rigid motion, this

work is not attempting to explain the operation of a presumed component of the human visual system for the interpretation of the non-rigid motion of such smoothly shaded surfaces.

1.4 Thesis Outline

Now that the topic of interest, the determination of three-dimensional velocities over deforming visible surfaces, has been introduced and motivated, the more specific topic addressed should be restated. This work considers how the changing intensity pattern induced by a smoothly deforming Lambertian surface illuminated by a distant point source is locally constrained by the three-dimensional motion of the surface. It then shows that the application of an additional constraint on the motion, (implemented in the form of a regulariser), enables certain psychologically significant classes of three-dimensional velocity field over the surface to be estimated veridically from the image intensity rate, the velocity field along the boundary and static information.

Chapter 2, the Background, introduces the mathematical notation relating to Differential Geometry which is required to understand the subsequent chapters, briefly summarises the mathematical technique of Regularisation, and reviews (with some minor novel results by way of comment) the most relevant work in the field of computer vision.

Chapter 3, on the Intensity Rate Constraint, determines the relationship between the image intensity change and the local surface deformation for a surface with Lambertian reflectivity (or a similar form), demonstrating the existence of a form of "aperture problem", the intensity rate being independent of the tangential components of surface velocity.

Chapter 4, on the Interpretation of Cylindrical Deformation, proposes a method by which the velocity field, including the tangential velocity, may be found (correctly in the case of a pure bending motion), for a reduced dimen-

sional form of the problem where the surface may be considered to be a curve undergoing deformation in a plane.

Chapter 5, on the Interpretation of Smooth Surface Deformation, extends the method of the previous chapter to the more general case of (potentially) doubly curved surfaces and illustrates the performance of the method with noise in the input data.

Chapter 6 summarises what has been found: the nature of the underlying constraint, the theoretical possibility of its application and the practical inapplicability with current technology due to the requirement for measurement accuracy. A suggested method for the interpretation of deforming three-dimensional curve motion is then proposed.

Chapter 2

Background

2.1 Introduction

This chapter consists of two major parts, the first introducing the mathematical notation and prerequisites for the rest of the thesis, and the second reviewing related computer vision work.

The first part covers two topics - Differential Geometry and Regularisation. These will both be treated (as will all mathematics in this thesis) from the perspective of engineering mathematics, *e.g.* it may be implicitly assumed where required that functions are *smooth* (*i.e.* have, at all points, derivatives of all orders) within the region of interest, having been previously segmented into such regions as required. The author hopes that the more mathematically refined reader will forgive this somewhat rough(!) approach.

The second part also covers two topics - Visual Motion and Shape from Shading.

2.2 Differential Geometry

Classically, differential geometry was the application of differential calculus to the study of the local properties of curves and surfaces in Euclidean space. It is this aspect of the field that is considered in this thesis, not the more abstract extension of the concepts as occurs in modern differential geometry.

The general material reviewed here, including that on deformation and differential invariants, may be found in greater detail in (Weatherburn, 1931; Weatherburn, 1930). Consequently any derivations will just be sketched out. The material on static shape may also be found in (doCarmo, 1976) using a similar notation. Some concepts will be illustrated with examples of relevance to later chapters. (A recent text in modern notation, with specific application to vision, is (Koenderink, 1990)).

2.2.1 Curves

A (smooth regular parameterised) three dimensional space curve ($\mathbf{r}(\alpha)$) is a mapping from $\Re$ into $\Re^3$, which has derivatives of all orders, the first being nowhere zero. It is convenient to parameterise a curve by its *arc-length* s, the arc-length from a to b being defined as:

$$s(\alpha) = \int_a^b \left| \frac{d\mathbf{r}}{d\alpha} \right| d\alpha \tag{2.1}$$

The unit *tangent* vector to the curve is defined by:

$$\mathbf{t} = \frac{d\mathbf{r}}{ds} \tag{2.2}$$

The unit *(principal) normal* vector to the curve $\mathbf{n}$ and the *(principal) curvature* κ are defined by:

$$\kappa\mathbf{n} = \frac{d\mathbf{t}}{ds} \tag{2.3}$$

(Note that the direction of $\mathbf{n}$ and the consequent sign of κ are arbitrary.)

The plane at $\mathbf{r}$ that contains $\mathbf{t}$ and $\mathbf{n}$ is termed the *osculating* plane, and the unit normal vector which is perpendicular to this plane is termed the *binormal* vector, defined as:

$$\mathbf{b} = \mathbf{t} \times \mathbf{n} \tag{2.4}$$

The formulae relating the derivatives of these three mutually perpendicular unit vectors are known as the Serret-Frenet formulae:

$$\frac{d\mathbf{t}}{ds} = \kappa\mathbf{n}$$

$$\begin{aligned} \frac{d\mathbf{n}}{ds} &= \tau\mathbf{b} - \kappa\mathbf{t} \\ \frac{d\mathbf{b}}{ds} &= -\tau\mathbf{n} \end{aligned} \tag{2.5}$$

Where τ is the *torsion* (which may take the opposite sign according to the preference of the author). While the curvature may be informally regarded as representing the extent to which a straight line would have to be bent in the plane to locally fit the (trace of the) curve, the torsion may be regarded as the extent to which it would have to be twisted. Thus, for planar curves the torsion is zero, the binormal vector being constant and perpendicular to the plane of the curve.

The case of the planar curve (in the $x - y$ plane) may be illustrated with the example of the circle, shown in Figure 2–1 given by

$$\mathbf{r}(\theta) = (r\,(\cos\theta - 1)\,, r\sin\theta)$$

where r is the radius, and $0 \leq \theta < 2\pi$.

This may be reparameterised in terms of the arc-length $s = r\theta$ as:

$$\mathbf{r}(\theta) = \left(r\left(\cos\frac{s}{r} - 1\right), r\sin\frac{s}{r}\right) \tag{2.6}$$

If the outward pointing normal is chosen, as shown in Figure 2–1, the (unit) tangent and normal vectors are then

$$\begin{aligned} \mathbf{t} &= (-\sin\theta, \cos\theta) \\ \mathbf{n} &= (\cos\theta, \sin\theta) \end{aligned}$$

(For convenience all vectors explicitly written out will be given as row vectors, with an implicit transpose symbol (^{t}) assumed). The binormal vector is, of course, perpendicular to the $x - y$ plane, (in the direction into the paper). Given these definitions, it is straightforward to check that the curvature κ is given by:

$$\kappa = -\frac{1}{r}$$

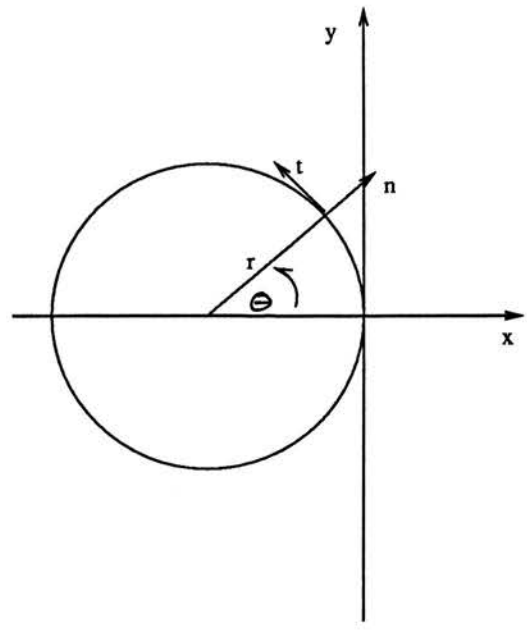

Figure 2–1: Circle Example

2.2.1.1 Curve Deformation

In the case of a curve undergoing an instantaneous motion such that the velocity at any point is $\mathbf{v}$ (a function of arc-length and time), the rate of extension per unit length, e, may be found by considering the change in length (δs) of an element $d\mathbf{r}$ after time δt. (The following is based on the proof on p.167 of (Weatherburn, 1930)). The new length is given by:

$$\sqrt{\left(d\mathbf{r} + \left(\frac{\partial \mathbf{v}}{\partial s}\delta s \delta t\right)\right)^2} = \sqrt{(d\mathbf{r})^2 + 2d\mathbf{r} \cdot \left(\frac{\partial \mathbf{v}}{\partial s}\delta s \delta t\right) + \left(\frac{\partial \mathbf{v}}{\partial s}\delta s \delta t\right)^2}$$

Ignoring the components of second order in δt yields:

$$\sqrt{(d\mathbf{r})^2 + 2d\mathbf{r} \cdot \left(\frac{\partial \mathbf{v}}{\partial s}\delta s \delta t\right)} = \delta s \sqrt{1 + 2\delta t \mathbf{t} \cdot \frac{\partial \mathbf{v}}{\partial s}}$$

Hence the rate of extension per unit length is:

$$e = \mathbf{t} \cdot \frac{\partial \mathbf{v}}{\partial s} \tag{2.7}$$

For a motion to be inextensible, e must be zero at all points on the curve, imposing the following condition upon the tangential and normal components of velocity (v_t and v_n respectively):

$$\frac{\partial v_t}{\partial s} = \kappa v_n \tag{2.8}$$

In the case of an inextensible motion, in which the arc-length is (by definition) independent of time, the order of differentiation with respect to time and arc-length may be inter-changed thus:

$$\frac{\partial \mathbf{v}}{\partial s} = \frac{\partial}{\partial s}\frac{d\mathbf{r}}{dt} = \frac{d}{dt}\frac{\partial \mathbf{r}}{\partial s} = \frac{d\mathbf{t}}{dt} \tag{2.9}$$

As the tangent vector is of constant (unit) length, the above represents a rotation of the tangent vector, the normal and bi-normal vectors undergoing a corresponding rotation. Following convention, ω will be used to denote the angular velocity, so that *e.g.*:

$$\frac{d\mathbf{t}}{dt} = \omega \times \mathbf{t} \tag{2.10}$$

Two example motions of an arc of the circle from the previous section will be used to illustrate a bending (*i.e.* unstretching) motion and a uniform expansion. The two forms of curve motion are firstly the bending motion:

$$\mathbf{r}(s,t) = \left((r+\dot{r}t)\left(\cos\left(\frac{s}{(r+\dot{r}t)}\right)-1\right),(r+\dot{r}t)\sin\left(\frac{s}{(r+\dot{r}t)}\right)\right)$$

and secondly the stretching motion:

$$\mathbf{r}(s,t) = \left(r\left(\cos\left(\frac{s+et}{r}\right)-1\right),r\sin\left(\frac{s+et}{r}\right)\right)$$

both considered only at the moment where $t = 0$. (In the second case, s is the arc-length only at $t = 0$). $\dot{r}$ and e are the rate of radius increase and the rate of extension respectively. The two types of motion are (roughly) illustrated in Figure 2–2, which shows the circle increasing in radius without stretching (whilst still passing through the origin), and Figure 2–3, which shows the circle stretching along itself.

The tangential and normal velocity components in the bending example are:

$$\begin{aligned} v_t &= \dot{r}(\sin\theta - \theta) \\ v_n &= \dot{r}(1 - \cos\theta) \end{aligned}$$

Hence

$$\frac{\partial v_t}{\partial s} = \frac{\dot{r}}{r}(\cos\theta - 1) = \kappa v_n$$

as required of an inextensible motion.

The normal velocity is zero in the stretching case, and the tangential velocity is:

$$v_t = re\theta$$

Hence the rate of extension is

$$\mathbf{t} \cdot \frac{\partial \mathbf{v}}{\partial s} = \mathbf{t} \cdot (e\mathbf{t} + \kappa re\theta\mathbf{n}) = e$$

as was intended by the definition of the motion.

2.2.2 Surfaces

A (smooth regular parameterised) three dimensional surface ($\mathbf{r}(\alpha, \beta)$) is a mapping from $\Re^2$ into $\Re^3$, which has partial derivatives of all orders, the vector cross product of the first order derivatives being non-zero, and has a continuous inverse at all points of the surface. The plane spanned by the first order derivatives ($\mathbf{r}_\alpha$ and $\mathbf{r}_\beta$) at a point on the surface is the *tangent plane* at that point and the unit vector perpendicular to the tangent plane

$$\mathbf{n} = \frac{\mathbf{r}_\alpha \times \mathbf{r}_\beta}{|\mathbf{r}_\alpha \times \mathbf{r}_\beta|} \tag{2.11}$$

is the surface normal. (Note that the direction of the normal, although arbitrary in the sense that a given surface in space may be parameterised in different ways, is fixed by the choice of parameterisation.)

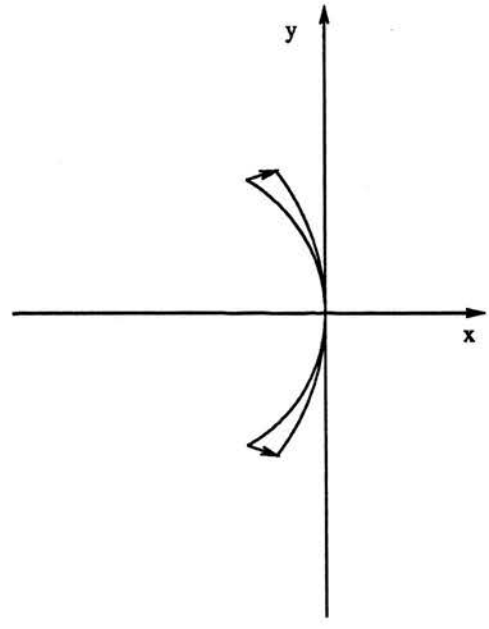

Figure 2–2: Bending Circle Example

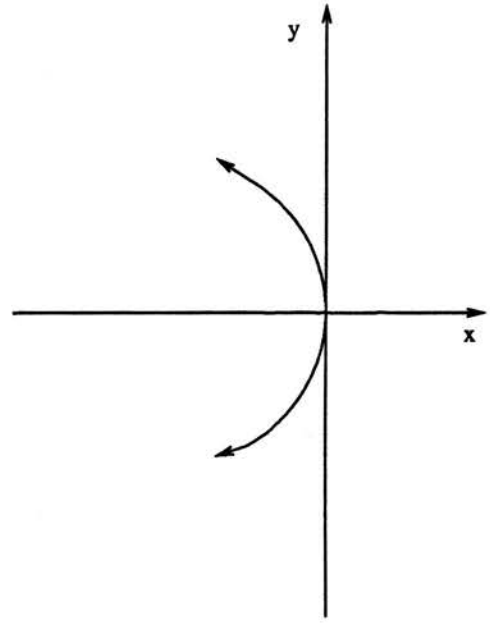

Figure 2–3: Stretching Circle Example

2.2.2.1 First Fundamental Form

The following expression, representing the square of the distance along the surface between two points separated by the infinitesimal increments $d\alpha$ and $d\beta$ in parameter space, is known as the *First Fundamental Form* of the surface:

$$E(d\alpha)^2 + 2Fd\alpha d\beta + G(d\beta)^2 \tag{2.12}$$

and the coefficients

$$\begin{aligned} E &= (\mathbf{r}_\alpha)^2 \\ F &= \mathbf{r}_\alpha \cdot \mathbf{r}_\beta \\ G &= (\mathbf{r}_\beta)^2 \end{aligned} \tag{2.13}$$

are known as the *First Order Magnitudes.* The quantity

$$A = \sqrt{EG - F^2} = |\mathbf{r}_\alpha \times \mathbf{r}_\beta| \tag{2.14}$$

represents the surface area per unit area of the parameter space, and is of sufficient significance to warrant its representation by a symbol. The symbol H is used in (Weatherburn, 1931; Weatherburn, 1930) and appears a natural progression from E,F and G, but unfortunately this clashes with the prevailing current conventional use of H in differential geometry; thus the symbol A shall be used hereafter, being a suitably mnemonic choice for an area quantity.

The first order magnitudes may be illustrated using two cases, the first a circular cylinder, based upon the planar circle used in the section on curves, and the second a torus.

The circular cylinder, illustrated in Figure 2–4, may be parameterised thus:

$$\mathbf{r}(\alpha, \beta) = \left(\alpha, r\sin\frac{\beta}{r}, r\left(\cos\frac{\beta}{r}\right)\right) \tag{2.15}$$

The torus, illustrated in Figure 2–5, may be parameterised thus:

$$\mathbf{r}(\alpha, \beta) = ((a + r\cos\beta)\sin\alpha, r\sin\beta, (a + r\cos\beta)\sin\alpha) \tag{2.16}$$

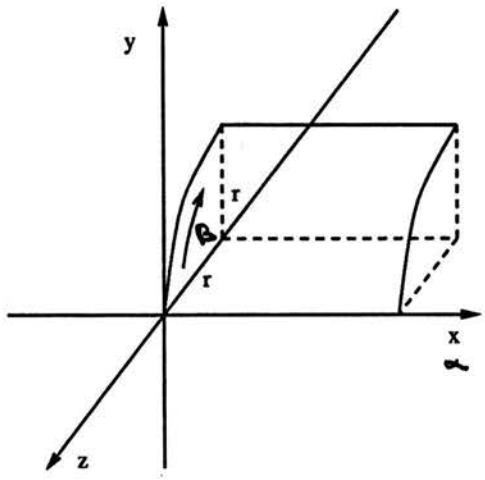

Figure 2–4: Circular Cylinder Example

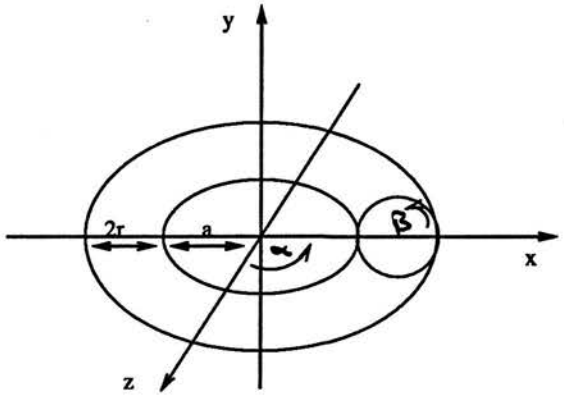

Figure 2–5: Torus Example

The first order magnitudes and A for the cylinder and torus are respectively:

$$\begin{aligned} E &= 1 \\ F &= 0 \\ G &= 1 \\ A &= 1 \end{aligned}$$

and

$$\begin{aligned} E &= (a + r\cos\beta)^2 \\ F &= 0 \\ G &= r^2 \\ A &= r(a + r\cos\beta) \end{aligned}$$

2.2.2.2 The Second Fundamental Form

The following expression, representing (twice) the length of the perpendicular from the tangent plane at a point on the surface to a point separated by $d\alpha$, $d\beta$ in parameter space is known as the *Second Fundamental Form* of the surface:

$$L(d\alpha)^2 + 2Md\alpha d\beta + N(d\beta)^2 \tag{2.17}$$

and the coefficients

$$\begin{aligned} L &= \mathbf{n} \cdot \mathbf{r}_{\alpha\alpha} \\ M &= \mathbf{n} \cdot \mathbf{r}_{\alpha\beta} \\ N &= \mathbf{n} \cdot \mathbf{r}_{\beta\beta} \end{aligned} \tag{2.18}$$

are known as the *Second Order Magnitudes.* Whereas the first fundamental form is an *intrinsic* quantity, *i.e.* is independent of the way the surface is embedded in $\Re^3$, the second is an *extrinsic* quantity depending upon (and expressing) that embedding.

Given the above definition, the orthogonality of the normal to $\mathbf{r}_\alpha$ and $\mathbf{r}_\beta$ may be used to derive expressions relating the first derivatives of $\mathbf{n}$ to the second order magnitudes *e.g.* differentiating $\mathbf{n} \cdot \mathbf{r}_\alpha = 0$ by α gives:

$$\mathbf{n}_\alpha \cdot \mathbf{r}_\alpha + \mathbf{n} \cdot \mathbf{r}_{\alpha\alpha} = 0$$

Thus

$$\mathbf{n}_\alpha \cdot \mathbf{r}_\alpha = -L \tag{2.19}$$

Similarly

$$\mathbf{n}_\beta \cdot \mathbf{r}_\beta = -N \tag{2.20}$$

The second order magnitudes of the circular cylinder and the torus of the previous section are respectively:

$$\begin{aligned} L &= 0 \\ M &= 0 \\ N &= \frac{-1}{r} \end{aligned}$$

and

$$\begin{aligned} L &= -(a + r\cos\beta)\cos\beta \\ M &= 0 \\ N &= -r \end{aligned}$$

2.2.2.3 Surface Curvatures

The *normal curvature* in a given direction at a point on a surface is defined in terms of the projection of the curvature of a surface curve tangential to that direction at the point:

$$\kappa_{(surface)} = \kappa_{(curve)} \left(\mathbf{n}_{(curve)} \cdot \mathbf{n}_{(surface)} \right)$$

The normal curvature is independent of the choice of such a curve for a specific direction but may vary with the direction - *umbilic* points, at which the normal

curvature is independent of direction will frequently be disregarded in what follows. The normal curvature reaches its maximum and minimum, known as the *principal curvatures* in the *principal directions* which are perpendicular.

Curves which are always tangential to a principal direction are known as *lines of curvature* - the coordinate curves of a parameterisation are lines of curvature if the coefficients F and M of the fundamental forms are everywhere zero. This constitutes an extremely convenient choice of parameterisation and **will hereafter be presumed** with little loss of generality, as (except in the neighbourhood of an umbilic point) a region of a surface can always be referred to its lines of curvature as parameters. This results in the principal curvatures being L/E in the direction $d\alpha = 0$ and N/G in the direction $d\beta = 0$. The parameterisations of the cylinder and torus already given were in terms of their lines of curvature.

Two related quantities frequently used in the characterisation of surfaces are the *mean curvature*, usually denoted H, the mean of the principal curvatures, and the *Gaussian curvature*, usually denoted K, the product of the principal curvatures. Whereas the sign of the mean curvature depends upon the choice of parameterisation, that of the Gaussian does not; thus the latter is of more significance in distinguishing between types of surfaces. However, in the application under consideration here, the sum of the principal curvatures (twice the mean curvature), denoted J is the most useful combination of the curvatures to have a symbol for, where

$$J = \frac{L}{E} + \frac{N}{G} \tag{2.21}$$

For the circular cylinder and the torus, J is given by the following expressions respectively:

$$J = \frac{-1}{r}$$

and

$$J = \frac{-1}{r} - \frac{\cos\beta}{(a + r\cos\beta)}$$

2.2.2.4 Differential Invariants

An operator ∇ may be defined (as in chapter XII of (Weatherburn, 1931)) for the purposes of considering the gradient of scalar functions defined over a surface and the divergence and curl of vector functions:

$$\nabla = \frac{1}{E}\mathbf{r}_\alpha \frac{\partial}{\partial \alpha} + \frac{1}{G}\mathbf{r}_\beta \frac{\partial}{\partial \beta} \tag{2.22}$$

When applied to the vector function

$$\mathbf{v} = v_\alpha \mathbf{r}_\alpha + v_\beta \mathbf{r}_\beta + v_n \mathbf{n} \tag{2.23}$$

the resulting divergence and curl may be expressed thus:

$$\nabla \cdot \mathbf{v} = \frac{1}{A}\left[\frac{\partial}{\partial \alpha}(Av_\alpha) + \frac{\partial}{\partial \beta}(Av_\beta)\right] - Jv_n \tag{2.24}$$

$$\nabla \times \mathbf{v} = \frac{1}{A}(Nv_\beta)\,\mathbf{r}_\alpha - \frac{1}{A}(Lv_\alpha)\,\mathbf{r}_\beta + \frac{1}{A}\left[\frac{\partial}{\partial \alpha}(Gv_\beta) - \frac{\partial}{\partial \beta}(Ev_\alpha)\right]\mathbf{n} + \nabla v_n \times \mathbf{n} \tag{2.25}$$

∇ can be seen to reduce to the well known form in the Cartesian $x - y$ plane:

$$\nabla = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y}$$

where $\mathbf{i}$ and $\mathbf{j}$ are the unit vectors in the x and y directions respectively. A three-dimensional form of ∇, ∇_3, may also be defined for application to functions defined over space parameterised by orthogonal coordinates α, β and γ:

$$\nabla_3 = \frac{1}{\mathbf{r}_\alpha \cdot \mathbf{r}_\alpha}\mathbf{r}_\alpha \frac{\partial}{\partial \alpha} + \frac{1}{\mathbf{r}_\beta \cdot \mathbf{r}_\beta}\mathbf{r}_\beta \frac{\partial}{\partial \beta} + \frac{1}{\mathbf{r}_\gamma \cdot \mathbf{r}_\gamma}\mathbf{r}_\gamma \frac{\partial}{\partial \gamma} \tag{2.26}$$

Again this reduces to the familiar form in Cartesian space, and the gradient of a scalar function Φ may also be expressed as

$$\nabla_3 \Phi = \frac{d\Phi}{d\mathbf{r}} \tag{2.27}$$

which will be used later.

2.2.2.5 Surface Deformation

In the case of a surface undergoing an instantaneous motion which may be represented as $\mathbf{v}$ (a function of α and β), the rate of area increase per unit of area may be found by considering the change in A at a point of the surface after an infinitesimal time increment δt.(The following proof is sketched from P.170 of (Weatherburn, 1930)). The new A is given by:

$$\sqrt{\left[\frac{\partial}{\partial\alpha}(\mathbf{r}+\mathbf{v}\delta t)\right]^2\left[\frac{\partial}{\partial\beta}(\mathbf{r}+\mathbf{v}\delta t)\right]^2}$$

$$=\sqrt{\left[\mathbf{r}_\alpha^2+2\mathbf{r}_\alpha\cdot\mathbf{v}_\alpha\delta t+(\mathbf{v}_\alpha\delta t)^2\right]\left[\mathbf{r}_\beta^2+2\mathbf{r}_\beta\cdot\mathbf{v}_\beta\delta t+(\mathbf{v}_\beta\delta t)^2\right]}$$

Ignoring quantities of second order in δt and substituting in the first order magnitudes yields:

$$\sqrt{EG+2\left[E\mathbf{r}_\beta\cdot\mathbf{v}_\beta\delta t+G\mathbf{r}_\alpha\cdot\mathbf{v}_\alpha\delta t\right]}=A\sqrt{1+2\nabla\cdot\mathbf{v}\delta t}$$

Hence the rate of area increase per unit area is (as should be expected):

$$\frac{dA}{dt}=\nabla\cdot\mathbf{v} \tag{2.28}$$

Similarly, the rate of change of the unit surface normal may be found by consideration of the normal at a point on the surface after an infinitesimal time δt. (The following proof is sketched from P.171–2 of (Weatherburn, 1930)). The new normal is given by:

$$\frac{1}{A(1+\nabla\cdot\mathbf{v}\delta t)}\left[(\mathbf{r}_\alpha+\mathbf{v}_\alpha\delta t)\times(\mathbf{r}_\beta+\mathbf{v}_\beta\delta t)\right]$$

Separating out the terms in δt in the numerator and ignoring those of second order gives:

$$\frac{1}{A(1+\nabla\cdot\mathbf{v}\delta t)}\left[(\mathbf{r}_\alpha\times\mathbf{r}_\beta)+(\mathbf{r}_\alpha\times\mathbf{v}_\beta-\mathbf{r}_\beta\times\mathbf{v}_\alpha)\,\delta t\right]$$

These terms may then be re-expressed in differential invariants of $\mathbf{v}$ to give:

$$\frac{1}{A(1+\nabla\cdot\mathbf{v}\delta t)}\left[(\mathbf{r}_\alpha\times\mathbf{r}_\beta)\,(1+\nabla\cdot\mathbf{v}\delta t)-(\mathbf{r}_\alpha\times\mathbf{r}_\beta)\times(\nabla\times\mathbf{v})\,\delta t\right]$$

Thus, ignoring terms of second order and above in δt, the rate of change of the unit surface normal is:

$$\frac{d\mathbf{n}}{dt} = (\nabla \times \mathbf{v}) \times \mathbf{n} \tag{2.29}$$

Example motions, based upon the circular cylinder and torus of the prior sections, will be used to illustrate these points. The first consists of the circular cylinder of Figure 2–4 bending in a manner similar to the curve of Figure 2–2, giving a motion which may be parameterised as:

$$\mathbf{r}(\alpha,\beta,t) = \left(\alpha, (r+\dot{r}t)\sin\left(\frac{\beta}{(r+\dot{r}t)}\right), (r+\dot{r}t)\left(\cos\left(\frac{\beta}{(r+\dot{r}t)}\right)-1\right)\right) \tag{2.30}$$

This yields the following components of surface velocity at time $t = 0$:

$$\begin{aligned} v_\alpha &= 0 \\ v_\beta &= \dot{r}(\sin\frac{\beta}{r} - \frac{\beta}{r}) \\ v_n &= \dot{r}(1-\cos\frac{\beta}{r}) \end{aligned}$$

The second motion, that of the torus expanding uniformly at a rate Θ may be parameterised as:

$$\mathbf{r}(\alpha,\beta,t)$$

$$= ((1+\Theta t)(a+r\cos\beta)\sin\alpha, (1+\Theta t)r\sin\beta\cos\omega t, (1+\Theta t)(a+r\cos\beta)\cos\alpha) \tag{2.31}$$

The corresponding components of surface velocity at time $t = 0$ are:

$$\begin{aligned} v_\alpha &= 0 \\ v_\beta &= -\frac{\Theta a}{r}\sin\beta \\ v_n &= \Theta(a\cos\beta + r) \end{aligned}$$

If the corresponding divergences are considered, that of the bending cylinder is straightforwardly zero, as may be seen by comparison of the expression for the divergence with that for the rate of extension of a deforming curve, and

the parameterisation of the bending circle and that of the bending circular cylinder. The divergence for the expanding torus is:

$$\nabla \cdot \mathbf{v} = \frac{A_\alpha}{A} v_\alpha + \frac{\partial v_\alpha}{\partial \alpha} + \frac{A_\beta}{A} v_\beta + \frac{\partial v_\beta}{\partial \beta} - J v_n$$

$$= 0 + 0 + \frac{-r^2 \sin\beta}{r(a + r\cos\beta)} \left(\frac{-\Theta a}{r} \sin\beta \right) - \frac{\Theta a}{r} \cos\beta$$

$$- \left(\frac{-1}{r} - \frac{\cos\beta}{(a + r\cos\beta)} \right) (\Theta(a\cos\beta + r))$$

which simplifies to 2Θ as should be expected – if lengths are expanded at the rate Θ, areas will expand at the rate 2Θ.

2.3 Regularisation

Regularisation is a method for determining a solution to an *ill-posed* mathematical problem, by constraining the solution to be, in some sense, regular or smooth. The best known work in this area is that of Tikhonov, see *e.g.* (Tikhonov & Arsenin, 1977), and it has been shown that many solutions to problems in low-level vision are examples of the method, see *e.g.* (Bertero et al, 1987), which also contains a fuller, more rigorous description of the method than will be presented here.

2.3.1 Ill-Posedness

The ill-posed problem is defined as one which is not well-posed, where the problem of determining z in the metric space Z from u in the metric space U, given the relationship:

$$z = R(u)$$

is ((Tikhonov & Arsenin, 1977)) *"well-posed on the pair of metric spaces" (Z,U)* if:

1. **The existence condition** For every $u \in U$ there exists a solution $z \in Z$, and

2. The uniqueness condition The solution is unique, and

3. The stability condition The problem is stable on the spaces (Z,U).

Conditions 1 and 2 can often be considered as characterising the mathematical determinacy of the problem, and 3 the physical determinacy and practicability of obtaining a numerical solution to a problem. In the last respect, it should be noted that regularisation methods can be applied to problems which are numerically ill-conditioned, though theoretically continuous.

A major category of ill-posed problems are *inverse problems* of the form:

$$Az = u$$

in which the problem of determining u from z is well-posed (and usually straightforward); however the problem of finding z given u is not.

A trivial example may be provided by considering the case where U and Z are both the space of real scalars $\Re$, and A the modulus operator. With the exception of the case where z is zero, the knowledge of the modulus of z does not uniquely specify z. This simple example may also be used to illustrate that a change in the definition of the required solution space Z of an ill-posed problem may create a well-posed one, albeit usually of little interest. Thus, in this case, redefining Z as the space of non-negative reals, creates a yet more trivial problem which is well-posed.

2.3.2 The Application of Smoothing Functionals to Ill-Posed Problems

The (ill-posed) problems to which regularisation methods are usually applied are those in which the solution z is not, as in the trivial example above, a scalar but a function of one or more variables sought over a range of these variables, *e.g.* the depth function may be sought over an image. The natural choice of metric is thus usually in the form of a functional.

The best known regularisation method is that of minimising (with respect to z) a functional of the form:

$$\| Az - u \|_U^2 + \lambda \Omega[z]$$

where the first term, a norm in the space U, represents the fidelity of the solution to the original data and the second its smoothness using the measure Ω, a stabilising functional chosen, usually from physical considerations of the original problem, to constrain the solution. λ is the *regularisation parameter* which controls the trade-off between fidelity and smoothness, of which Tikhonov comments in (Tikhonov & Arsenin, 1977)

> In the study of particular problems it is usually difficult to actually find the regularisation parameter ...

However, the same text does suggest a number of methods for deriving a suitable value of λ. More recent suggestions include the *Min-Max Principle* proposed in (Gennert & Yuille, 1988), in which λ is chosen to maximise the minimum obtained with respect to z (to avoid the possibility that either term predominates), and (for discrete data) the method of *Generalised Cross Validation* (Wahba, 1980) in which the parameter is chosen by consideration, for each data item, of the error in a prediction of that item from a data set which excludes it. Note that as $\lambda \to 0$, the method becomes equivalent to minimising Ω subject to the original equation as a constraint.

2.3.2.1 Tikhonov Stabilisers

Many applications of regularisation use a particular class of stabilising functional - the *Tikhonov Stabiliser*. In the case where $z(x)$ is sought on the interval $[a, b]$ and Z is the (Sobolev) space of functions with square-integrable generalised derivatives up to pth order with the metric:

$$\left[\int_a^b \sum_{r=0}^{p} q_r(x) \left(\frac{d^r z}{dx^r} \right)^2 dx \right]^{\frac{1}{2}}$$

where the $q_r(x)$ are given non-negative continuous functions, $q_p(x)$ a given positive continuous function, a regularising functional of the following form may be used:

$$\Omega[z] = \int_a^b \sum_{r=0}^{p} q_r(x) \left(\frac{d^r z}{dx^r}\right)^2 dx$$

If all $q_r(x)$ are constants, the functional is a *Tikhonov stabiliser* - one of the most common in low-level vision applications.

2.4 Visual Motion Analysis

The extensive research on the topic of visual motion analysis may be grouped (or divided) in more than one manner. The primary criterion which will be used here is whether the method is based upon the assumption that the problem is one of analysing a discrete image pair (or sequence), or is based upon the assumption that the problem is one of analysing a continuously changing image. The indirect phrasing "based upon the assumption that the problem is" is used because in practice the methods for analysis of continuous change are applied to discrete image pairs (or sequences) with any required temporal derivatives being calculated from the inter-image differences. As this thesis follows the continuous approach, work in this area will be described first and at greater length.

2.4.1 Continuous Motion Analysis

The problem of calculating the environmental surface structure and motion from the perceived changing intensity image has classically been addressed in two parts. The first has been the determination of the *optic flow field* - the two-dimensional velocity field in the image induced by the projection of the three-dimensional velocity field of environmental motion relative to the observer. The second has been the derivation of the three-dimensional environmental velocity field and surface structure from a given optic flow field - usually under

the assumption that the entire environment is rigid or that the image flow field may be segmented into regions corresponding to rigid parts of the environment.

Thus, the work is reviewed here in three categories, the first two being the two parts of the classical approach, and the third part is that work which does not completely follow the classical approach, usually linking the three-dimensional calculation more directly to image quantities.

2.4.1.1 Optic Flow Field Calculation

The calculation of the veridical optic flow field for a changing image is not straightforward, and requires assumptions to be made about the relationship between the image and the viewed scene. The most common assumption to be made is that the optic flow is equivalent to the *image flow* - the two-dimensional velocity field which maps the image intensity pattern at one time instant on to that at the succeeding time instant. This assumption holds if the image intensity value for the (changing) image point corresponding to a point in the world remains constant. This is true for a number of cases *e.g.* the case of observer motion parallel to the image plane while viewing a fixed Lambertian scene under orthographic projection with unchanging illumination, but is not universally true *e.g.* the motion of light patterns along a discotheque wall are not usually indicative of the rate of motion of the wall! Some authors, *e.g.* (Nagel, 1987) use the term "optical flow" for what is termed here "image flow", on the grounds that in a practical application the true two-dimensional velocity field in the image induced by the projection of the three-dimensional velocity field of environmental motion relative to the observer is unknown and hence not a useful quantity to consider; however, it seems worth retaining the distinction to emphasise the potential gap between the veridical image flow field and the (desired) optic flow field.

The *motion constraint equation* which underlies most methods of flow recovery (see (Nagel, 1987) for a discussion of the relationship among many such methods only the more relevant of which will be dealt with here) will be con-

sidered before specific methods for recovery of flow over image regions and on image edges - corners will be omitted as they are not as relevant and the prior reference should suffice.

The Motion Constraint Equation The *motion constraint equation* (Horn & Schunck, 1981), is a first-order Taylor series expansion of the image intensity considered as a function of time and space $I(x, y, t)$. This yields the following relationship between the local image velocity and local spatial and temporal intensity derivatives:

$$I_x \frac{dx}{dt} + I_y \frac{dy}{dt} + I_t = 0 \tag{2.32}$$

where subscripts are used to denote partial differentiation. This may alternatively be expressed (using the conventional form of ∇ in the Cartesian plane and letting $\mathbf{v}$ be the two-dimensional image velocity) as:

$$\nabla I \cdot \mathbf{v} + I_t = 0 \tag{2.33}$$

The determination of image flow from the motion constraint equation is hindered by the *aperture problem* - the velocity along a direction of zero intensity gradient is locally irrecoverable. This is intuitively to be expected, as if one considers image motion directly along such a direction, it will be unobservable, as image points will move to points of equal intensity. This causes methods for flow recovery at points other than intensity extrema to require the imposition of further constraints upon the solution.

Image Flow Recovery Over a Region As the image flow is generally locally irrecoverable due to the aperture problem, the problem of image flow recovery is *ill-posed* and a further constraint must be imposed to enable flow recovery. Such a constraint should be motivated by knowledge of the form of the image motion. For example, in the hypothetical case of observer motion mentioned above, in which image flow and optic flow are equivalent, if it were known that the observer motion was a pure translation without rotation, the x and y velocity components would be independent of image position; thus one

could impose this as a constraint upon the solution. Hence, unless the direction of the intensity gradient was everywhere constant, the image flow could be determined (ignoring issues of noise). However, the image velocity will not in general be independent of position, so more general forms of constraint should be considered.

Horn and Schunck (1981) applied a smoothness constraint to the problem, (apparently unaware that this was a regularisation method), finding the velocity field which minimised the following integral over the region of interest:

$$\int\int \left[I_x \frac{dx}{dt} + I_y \frac{dy}{dt} + I_t\right]^2 + \alpha^2 \left[\left(\frac{dx}{dt}_x\right)^2 + \left(\frac{dx}{dt}_y\right)^2 + \left(\frac{dy}{dt}_x\right)^2 + \left(\frac{dy}{dt}_y\right)^2\right] dxdy \tag{2.34}$$

where α^2 was a positive constant and the subscript notation is for derivatives. This was formally minimised by the application of the Calculus of Variations and the resulting Euler-Lagrange equations discretised and solved iteratively; for results see the above reference.

This method was extended in (Cornelius & Kanade, 1983), in which the velocity smoothness constraint was only applied within regions, across the boundaries of which the velocity was permitted to vary without penalty, (a now common presumption when applying smoothness criteria). Further, rather than assume that the perceived intensity at the changing image point corresponding to a world point remained constant, the less restrictive assumption that the intensity change, $\frac{dI}{dt}$, should vary smoothly over the image (like the velocities) was applied, by adding a further smoothness term of the form

$$\left[\frac{\partial}{\partial x}\left(\frac{dI}{dt}\right)\right]^2 + \left[\frac{\partial}{\partial y}\left(\frac{dI}{dt}\right)\right]^2$$

and the motion constraint equation modified to allow for the non-zero $\frac{dI}{dt}$.

Image Flow Recovery Over a Curve The first stage in the analysis of raw intensity images in many vision applications is the detection of *edge* pixels (edgels), pixels at which the (smoothed) image intensity gradient is locally maximal and the linking of such edgels into edge curves. This approach is

motivated by the fact that such intensity edges frequently correspond to edges of objects in the scene and hence are salient features, not just artefacts of the imaging process. It is also known from physiological studies that animal retina contain cells which perform a process which may be regarded as edge detection. For these reasons, and additionally the fact that the motion constraint equation is better conditioned when the gradient term is large, a number of researchers have considered the problem of flow field recovery along a curve.

The best known such work is (Hildreth, 1984). In this, it was suggested that the image was first convolved with the Laplacian of a Gaussian operator, to detect the edges as zero-crossings of the resulting image, to which the motion constraint equation could be applied to yield estimates of the local image velocity normal to the contour. The locally immeasurable component of velocity tangential to the contour was then determined from the locally measurable component normal to the contour, by the application of a smoothing functional, (again, like (Horn & Schunck, 1981), apparently unaware that this was an example of regularisation). The measurement of (non-)smoothness used was:

$$\int \left(\frac{\partial \mathbf{v}}{\partial s}\right)^2 ds \tag{2.35}$$

where (as in the differential geometry section), $\mathbf{v}$ is the velocity vector (in this case two-dimensional), and s is the arc-length.

Yuille (1985) analysed Hildreth's method for the case in which the normal velocity is assumed to be perfectly known, hence the tangential velocity is found by minimising the above subject to the normal velocity as a constraint. It was shown that the correct motion field will only be found if the true motion satisfies

$$\frac{\partial^2 \mathbf{v}}{\partial s^2} \cdot \mathbf{t} = 0 \tag{2.36}$$

on the contour, where (as previously) $\mathbf{t}$ is the (unit) tangent vector to the contour.

Thus for example, rigid translation or uniform expansion of any image, and rotation of images consisting of straight lines are cases where the correct field will be found. By consideration of the equations of motion for an inextensible curve (see section 2.2.1), it may be seen that within *this* class of motions, rigid translation of an arbitrary contour, and rotation of straight lines are the *only* cases which are correctly solved, since for an inextensible curve:

$$\frac{\partial \mathbf{v}}{\partial s} = \frac{d\mathbf{t}}{dt} = \omega \times \mathbf{t}$$

thus:

$$\frac{\partial^2 \mathbf{v}}{\partial s^2} = \frac{\partial \omega}{\partial s} \times \mathbf{t} + \omega \times \frac{\partial \mathbf{t}}{\partial s}$$

hence from the Frenet relations (2.5):

$$\frac{\partial^2 \mathbf{v}}{\partial s^2} \cdot \mathbf{t} = \kappa \left(\omega \times \mathbf{n}\right) \cdot \mathbf{t}$$

Contour inextensibility has been explicitly used as a criterion for determining the full motion from the measured normal velocity in (D'Haeyer, 1986), which extended(!) this to determine the motion under a (square of) stretch-minimising assumption. The stretch measure used was:

$$\int \left(\frac{\partial \mathbf{v}}{\partial s} \cdot \mathbf{t}\right)^2 ds$$

Interestingly, contour inextensibility is implicitly assumed in the working in (Brady, 1987), in the following form:

$$\frac{\partial \mathbf{v}}{\partial s} = \frac{d\mathbf{t}}{dt}$$

This assumption enables the first arc-length derivative of the motion constraint equation to be simplified from:

$$\frac{\partial \nabla I}{\partial s} \cdot \mathbf{v} + \nabla I \cdot \frac{\partial \mathbf{v}}{\partial s} + \frac{\partial I_t}{\partial s} = 0 \tag{2.37}$$

to:

$$\frac{\partial \nabla I}{\partial s} \cdot \mathbf{v} = 0 \tag{2.38}$$

(For details of the working see the original paper.)

In (Gong, 1988), it was proposed that the following constraint be used (in the paper, the derivation of the constraint appears flawed, due to an erroneous Taylor series expansion):

$$\frac{\partial \nabla I}{\partial s} \cdot \frac{\partial \mathbf{v}}{\partial s} = 0$$

Under the assumption of inextensibility, this may be combined with 2.37, to yield, as per (Gong, 1988):

$$\frac{\partial \|\nabla I\|}{\partial s} \left(\frac{\partial \nabla I}{\partial s} \cdot \mathbf{v} + \frac{\partial \nabla I}{\partial t} \cdot \mathbf{t} \right) = 0$$

Taking $\frac{\partial \|\nabla I\|}{\partial s}$ to be non-zero, yields Gong's "curve motion constraint equation" (c.m.c.e.), which is equivalent to assuming $\nabla I \cdot \frac{\partial \mathbf{v}}{\partial s}$ is zero in 2.37. It is of interest to note that this assumption, which may be re-expressed as 2.40 below for non-zero intensity gradient, is of similar form to that of inextensibility (2.39 below), being a (different) constraint upon the arc-length derivative:

$$\frac{\partial \mathbf{v}}{\partial s} \cdot \mathbf{t} = 0 \tag{2.39}$$

$$\frac{\partial \mathbf{v}}{\partial s} \cdot \mathbf{n} = 0 \tag{2.40}$$

Clearly, the only velocity field which satisfies both is one of pure translation - perhaps the most common test case for such methods. Gong has demonstrated test results of a method combining Hildreth's smoothness criterion with the determination (where possible) of tangential velocity from his curve motion constraint equation (c.m.c.e). It is of interest to note that again a pure translation field will satisfy both 2.40 which underpins the c.m.c.e., and 2.36, which implies being veridically found under the smoothness assumption from the normal velocity along the contour. Further, a uniform contour extension satisfying 2.36 also satisfies 2.40, as for a uniform extension:

$$\frac{\partial}{\partial s} \left(\frac{\partial \mathbf{v}}{\partial s} \cdot \mathbf{t} \right) = 0$$

hence

$$\frac{\partial^2 \mathbf{v}}{\partial s^2} \cdot \mathbf{t} + \kappa \frac{\partial \mathbf{v}}{\partial s} \cdot \mathbf{n} = 0$$

Thus, it is apparent that the basic optic flow constraint, which is derived from the assumption that the image intensity corresponding to a point in the world remains constant, may be combined with a wide variety of assumptions relating to contour motion to yield the correct velocity field in the case of pure translation, (assuming ideal data). However, the choice of additional assumption(s) is critical to the correct determination of other types of motion, such as rotation or contour extension.

2.4.1.2 Optic Flow Interpretation

If the optic flow can be estimated, the resulting two-dimensional velocity field may then be interpreted in terms of the environmental structure and the relative motion between observer and environment which could have caused the flow field. This again requires assumptions to be made about the nature of the environment and a knowledge of the projective geometry involved in the image formation. The most common form of projection used in optic flow interpretation is perspective and the research area has been summarised as "fundamentally a study of the time-varying projective geometry of surfaces moving through space" (Waxman & Wohn, 1986). This section will summarise the general approach taken by (Subbarao, 1988) (to which the reader is referred for more detail of the methods and results) and others to the standard problem of interpreting optic flow assuming the environment is rigid - a powerful constraint upon the solution. The notation here will follow (Subbarao, 1988), which in turn follows the well-known earlier work of (Longuet-Higgins & Prazdny, 1980). Note that the conventions adopted for this section are not applicable elsewhere.

Consider, as in Figure 2–6, a point P on a smooth surface being viewed under perspective projection through the origin, with the image plane on the Z-axis (in front of the point of projection for mathematical convenience). Let the surface be at rest and the coordinate frame $OXYZ$ be in motion with a translational velocity and a rotational velocity. The motion of P in $OXYZ$ may then be found in terms of its position and the coordinate frame motion.

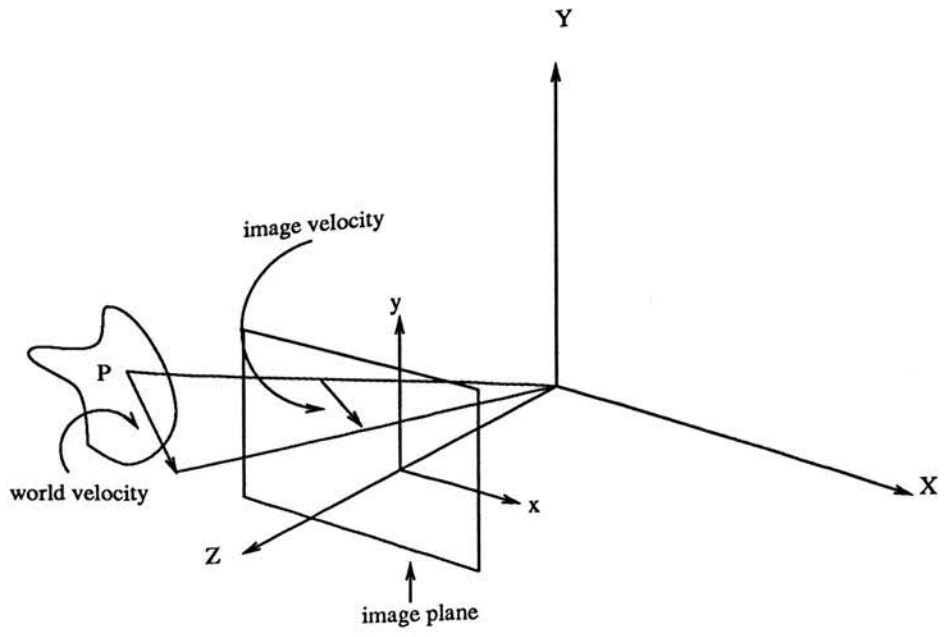

Figure 2–6: Moving coordinate frame used in motion analysis

Hence the motion in the image plane of the image point corresponding to P may be derived. The problem of flow interpretation is to interpret this image plane motion.

This is accomplished by considering a second order Taylor series of the surface shape in the form $Z(X,Y)$, and expanding the expression for the image velocity by substituting for Z. The result makes clear the inherent scale ambiguity as the translational velocity and distance along the Z-axis to the object always appear in ratio, hence their absolute value can never be recovered, only the ratio. As the surface is presumed to be smooth, and the projection continuous, the second order Taylor series of the image velocity may now be considered. The parameters of these series, which may be measured in the image, are found in terms of the unknowns – the scaled translational velocities, the angular velocities and the terms of the series expansion for Z. These equations may then be manipulated to yield (sometimes ambiguous) solutions for the unknowns in terms of the measurable image parameters.

Subbarao (1988) extended the above approach to cater for non-rigid motion, by considering the case in which the velocity of P is also considered to be a function of position (and time) and is thus itself approximated by a Taylor series expansion up to requisite order. In principle, it then appears that the optic flow may be approximated by a Taylor series expansion (potentially as a function of time as well as position) up to sufficiently high order to have sufficient equations for the number of unknowns. In practice the imposition of constraints, such as the assumption of rigidity above, is necessary as the general problem is underconstrained.

Koenderink and van Doorn (1986) also considered the case of point motion in continuous time, under perspective projection, but the points were assumed to be on a surface undergoing a bending (isometric) motion. This was approximated, using a difference geometry approach, by hinged planar facets, and the constraints on the motion at a single vertex considered. It was shown that partial shape information (with a relief ambiguity) may be recovered from the point motions and derived estimates of differential invariants of the motion field.

2.4.1.3 Other Methods for Continuous Motion Interpretation

The "classical" approach reviewed above, in which the task of interpreting continuous image change is split into the separate tasks of interpreting the intensity change as a two-dimensional velocity field in the image and then interpreting this in terms of the environmental three-dimensional velocity field, is an appropriate task decomposition in the cases where the intermediate representation, the image flow estimated from the first part, the optic flow into the second, is (barring issues of noise) the same item. Unfortunately, as has been shown above, even in those cases where the intensity at the changing image point corresponding to a world point does remain constant and hence veridical image flow is optic flow, the image flow is not necessarily found correctly (in the absence of noise) due to the aperture problem. Thus, alternative approaches

have been suggested to exploit the same constraints – the rigidity assumption and the motion constraint equation.

Buxton and Williams (1986) followed the classical split of determining image velocity prior to determining world velocities, but used only the known component of the image velocity, that in the direction of the intensity gradient, thus neatly side-stepping the aperture problem. This was accomplished by making an additional assumption of planarity to further constrain the solution, using a method developed in (Longuet-Higgins, 1984).

Negahdaripour and Horn (1986) used a (conceptually) one-step method to solve the problem of recovering the motion of planar surfaces from the changing intensity pattern, formulating the recovery of motion and surface orientation as a regularisation problem.

The motion work perhaps most related to that of this thesis is (Shulman & Aloimonos, 1988), which proposed a general regularisation framework for the determination of (non-)rigid motion from a changing intensity pattern. A number of previous techniques may be regarded as special cases of this framework and the case of the Lambertian surface, which is addressed in this thesis, was mentioned. However, although the paper "proposes a class of algorithms rather than a particular one", the approach of this thesis, which was developed independently, is quite distinct in its use of a coordinate system set in the surface. This enables the form of the constraint relating the image intensity change to the world velocity to be clarified in a novel manner which is not possible in the conventional Cartesian coordinate system employed in (Shulman & Aloimonos, 1988). The approach suggested in that paper (but apparently unimplemented) is, however, more general and in such respects as suggesting the application of "linear feature functionals" for smoothing, and the use of a multi-grid method of solution is hypothetically superior to the actual implementation in this thesis, which was developed only as a basic test of the theories suggested and has not been optimised for noise immunity or execution speed.

2.4.2 Discrete Motion Analysis

The analysis of discrete images in which the change between images is too large to permit the application of the differential based methods of continuous motion methods has also largely been based upon a two step approach. The images are presumed initially to consist of point tokens, which are corresponded between the images as a first stage, and the two dimensional point motions between the images then interpreted as a three dimensional object motion in the world in the second stage. As in the case of continuous motion, rigidity is frequently used to constrain the solution.

Ullman (1979) followed this approach, originating the application of the "rigidity assumption" and demonstrating that three dimensional structure may be recovered (up to a reflection ambiguity) in the case where four points are seen and corresponded among three orthographic views. When two views are considered, this may be regarded as equivalent to the stereo problem; *e.g.* Longuet-Higgins (1981) demonstrated a solution for the case of 8 points in 2 perspective views.

Ullman (1984) applied the two stage approach to the recovery of (non-)rigid structure (and hence motion) in a discrete time sequence of images under orthographic projection, containing corresponded tokens. For each new image a 3-D structure (initially flat) was hypothesised based on an optimisation process which minimises the sum of the squares of changes in 3-D inter-token distances between images. When this method works for a rigid structure the hypotheses converge on the true structure or its reflection about the image plane – equally valid solutions under orthographic projection. The method admits the possibility of calculating non-rigid structures in motion, but even for rigid structures the rate of convergence is highly dependent on the amount by which the viewpoint changes between images. Problems are of course to be expected when the sampling rate is inadequate relative to the motion, but less acceptable may be the instability in the limiting continuous time case, as illustrated in a paper based on Ullman's method (Grzywacz & Hildreth, 1987).

Aloimonos and Rigoutsos (1986) retained the rigidity assumption but did not presume that individual points could be corresponded. Instead, within a binocular framework, the motion between views of a set of points was recovered by consideration of the equations for the centroid and moments.

Chen and Penna (1986) considered the case of two corresponded intensity images under perspective projection of the same surface, which had undergone an isometric transformation between images. A Taylor series expansion for the 3-D surface motion around any point was then recovered from the correspondence, the equations of the projection, and surface shape information (nominally recovered by three source photometric stereo which will be discussed later).

Terzopolous *et al.* (1987) considered a discrete time series of stereo image pairs of a deforming object. The occluding contours were used to constrain the shape of a symmetry-seeking deformable model (with motion damping); thus instantaneous shape and non-rigid motion were recovered. It was intended to extend the method to use other constraints, such as reflectivity.

2.5 Shape From Shading

The shading of an object constitutes a cue as to its shape, as is demonstrated by the use of shading algorithms to create realism in the field of computer graphics; thus the interpretation of image shading as object shape has been studied in computer vision. Although this thesis rests upon the assumption that such a method may be used to determine the static shape of an object whose motion is of interest, it adds nothing in the field of such shape determination, so this aspect will be reviewed extremely tersely. (Horn, 1986) contains an excellent and lengthier introduction to the field.

The image intensity corresponding to a point in the world depends upon the illumination, the reflectivity of the point and the projection. In the general case where nothing is known *a priori*, nothing can be concluded about the

shape of the object. Methods for shape recovery generally presume that the projection is known and, typically, that the form of the *reflectance map*, the image intensity as a function of surface orientation is known.

Ikeuchi and Horn (1981) suggested a method for surface shape recovery from a knowledge of the occluding normals and the reflectance map by the method of characteristic strip expansion. Blake *et al.* (1985) investigated the recovery of surface shape and position from a knowledge of the 3D bounding curve (as could be found by stereo) and the reflectance map – information which the major part of this thesis assumes available for such a determination of the static shape.

The most common form of reflectivity presumed by methods is *Lambertian*, (perfectly matt) in which light is scattered equally in all directions from the surface. When a surface of constant Lambertian reflectivity is illuminated by a point light source at infinity (in direction $\mathbf{l}$) and viewed under orthographic projection, the perceived image intensity for a point with normal vector $\mathbf{n}$ is:

$$I = \rho \mathbf{l} \cdot \mathbf{n}$$

where ρ is the *albedo*, indicating what proportion of incident light is reflected by the surface.

Photometric Stereo, originated in (Woodham, 1978) and analysed in (Horn et al, 1978), is a method of using images taken with different known light source directions to recover the surface shape. If the albedo is known, the shape may be found using two light sources. If not, the albedo and shape may be found with the use of three sources.

Wolff (1989) proposed a variation upon Photometric Stereo, in which the orientation of the light source was changed continuously, rather than discretely, and hence the change in image intensity values may be regarded as a "photometric flow field", the value of which, in conjunction with the corresponding light source motion, may be used to derive local surface shape. This method which involved analysing the changing intensity pattern due to a light source motion over a fixed Lambertian rigid object, is perhaps the method in this

section most closely related to the main topic of this thesis – the analysis of a changing intensity pattern due to a (non-)rigid Lambertian object motion under fixed lighting.

Pentland (1984) combined the assumption of Lambertian reflectivity with an assumed known statistical distribution of surface curvatures and an assumption regarding the surface type to find the tilt and a maximum likelihood estimate of the slant.

Brooks and Horn (1985) used the application of smoothing functionals to recover an estimate of surface orientation, source vector and surface albedo by a minimisation over the surface.

2.6 Conclusions

The notation and major relevant concepts in the differential geometry of curves and surfaces and the method of regularisation, which will be used in subsequent chapters, have been introduced. The classical approach to the interpretation of changing intensity images as three dimensional surfaces in motion has been described. Some previous methods for image flow estimation over a contour have been analysed in terms of their capability on unstretching curve motion (or their assumption of such motion). The split between the derivation of image flow in the image plane and the interpretation of optic flow, to which the image flow is presumed to correspond, has also been made clear. Finally, the standard form of shading presumed in the interpretation of *shape from shading*, Lambertian shading, as will be presumed in this thesis, has also been covered.

Chapter 3

The Intensity Rate Constraint

3.1 Introduction

This chapter shows that for a deforming Lambertian surface of constant uniform albedo (viewed orthographically and illuminated by an infinitely distant point source), the local image intensity rate is independent of the tangential components of surface velocity, depending only on the normal component of surface velocity, the instantaneous shape and the illumination. The derivation of this intuitively obvious, but (as far as the author is aware) hitherto unproved fact is simplified by the use of a coordinate system set in the surface, rather than the Cartesian system customarily employed in the visual analysis of motion. This coordinate system is retained in subsequent chapters to emphasise the form of the constraint underlying the interpretation of the intensity rate, although a transformation into a Cartesian system could be preferable in a practical application.

3.2 Problem Definition

It is desired to derive an expression for the temporal intensity derivative at a point in a changing intensity image, which is the orthographic projection along the z-axis of an arbitrarily smoothly deforming smooth surface of constant uniform albedo Lambertian reflectance, illuminated by a point light source at infinity. (See Figure 3–1).

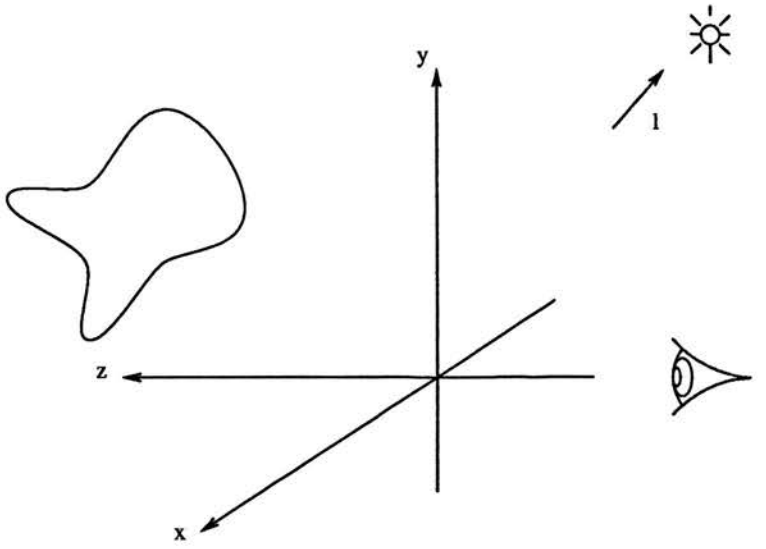

Figure 3–1: Diagram of viewing and illumination configuration

3.3 Static Image Intensity

For a Lambertian surface, at any instant, the image intensity corresponding to a point on the surface, will be:

$$I = \rho \mathbf{l} \cdot \mathbf{n}$$

where

- I is the image intensity (assuming appropriate sensor calibration)
- ρ is the surface albedo, hereafter absorbed into $\mathbf{l}$ for convenience
- $\mathbf{l}$ is the light source vector
- $\mathbf{n}$ is the unit normal vector on the surface

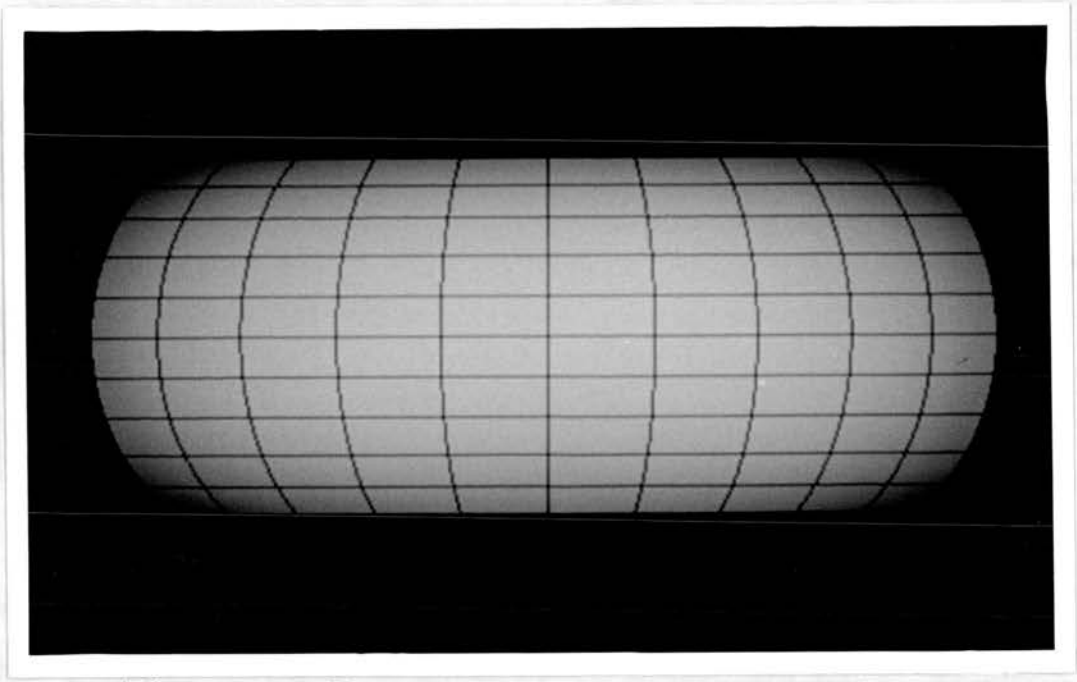

Figure 3–2: Lines of curvature on toroidal surface patch

3.4 Surface Coordinate System

The derivation will presume that the surface has been parameterised (in the form $\mathbf{r}(\alpha, \beta)$) by its lines of curvature and that all relevant quantities such as image intensity may be considered as functions of α and β, the parameters of the *visible* point corresponding to the image point. An example of this is illustrated in Figures 3–2 and 3–3, where part of a uniform albedo toroidal surface (illuminated by a point source behind the viewer) is parameterised by its lines of curvature (as added to the $x - y$ plane image), and the intensity as a function of these parameters illustrated below (with the lines of curvature now forming a rectangular grid).

Figure 3–3: Intensity displayed as a function of lines of curvature parameters

3.5 Derivation

Given the previous expression for the image intensity at a point, knowing that the light source is fixed, and presuming the surface reflectivity to be unchanged in motion, the change in intensity at the *changing* image point corresponding to a given point on the surface is :

$$\frac{dI}{dt} = \mathbf{l} \cdot \left(\frac{d\mathbf{n}}{dt} \right)$$

Substituting for the change in the unit normal vector gives:

$$\frac{dI}{dt} = \mathbf{l} \cdot ((\nabla \times \mathbf{v}) \times \mathbf{n})$$

where the vector fields $\mathbf{v}$ (translational velocity) and $\mathbf{n}$ are defined on the surface.

Considering the full and partial first-order derivatives of image intensity:

$$\frac{dI}{dt} = \dot{I} + (\nabla_3 I \cdot \mathbf{v})$$

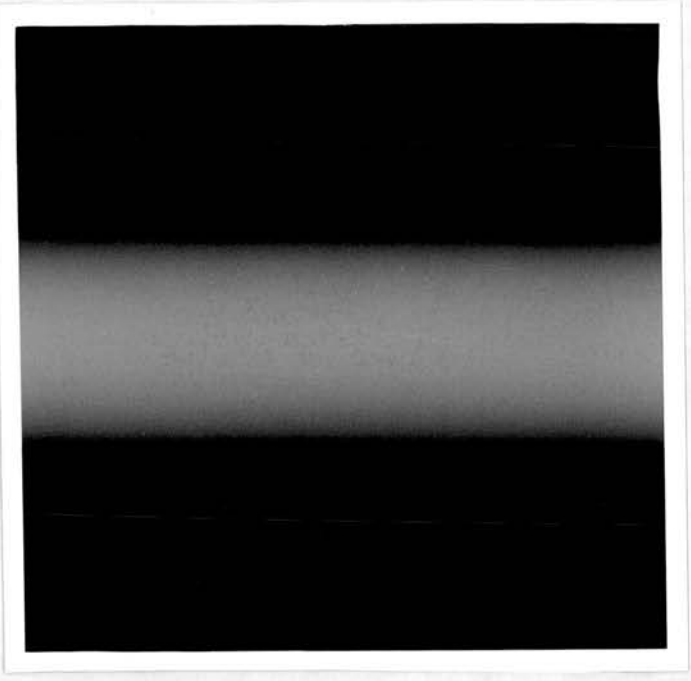

Figure 3–4: z-y Plane Section through "3-D" intensity function

Hence

$$\dot{I} + (\nabla_3 I \cdot \mathbf{v}) - \frac{dI}{dt} = 0 \tag{3.1}$$

where

- $\dot{I}$ is the intensity rate at a fixed image point
- $\nabla_3 I$ is the 3-D gradient of the intensity.

The latter concept may seem rather odd given that the intensity image is two-dimensional! It should be regarded as expressing the projection. In this case, (orthographic projection in the z direction), movement in the z direction will leave the image unaffected as the intensity gradient is zero in this direction. In the case of perspective projection, contraction towards, or expansion from, the point of projection is special in that a surface element remains in correspondence with an image plane element. A cross-section of the conceptual three-dimensional intensity resulting from the toroidal surface previously illustrated is shown in Figure 3–4. The terms involving tangential components of velocity in 3.1 may be eliminated by considering the expansions of the terms, using the basis $(\mathbf{r}_\alpha, \mathbf{r}_\beta, \mathbf{n})$:

$$\mathbf{v} = v_\alpha \mathbf{r}_\alpha + v_\beta \mathbf{r}_\beta + v_n \mathbf{n}$$

$$\mathbf{l} = l_\alpha \mathbf{r}_\alpha + l_\beta \mathbf{r}_\beta + l_n \mathbf{n}$$

Thus:

$$\nabla_3 I = \frac{1}{E}\frac{\partial}{\partial \alpha}(\mathbf{l}\cdot\mathbf{n})\,\mathbf{r}_\alpha + \frac{1}{G}\frac{\partial}{\partial \beta}(\mathbf{l}\cdot\mathbf{n})\,\mathbf{r}_\beta + I_n\mathbf{n}$$

$$= \frac{1}{E}(\mathbf{l}\cdot\mathbf{n}_\alpha)\mathbf{r}_\alpha + \frac{1}{G}(\mathbf{l}\cdot\mathbf{n}_\beta)\mathbf{r}_\beta + I_n\mathbf{n}$$

$$= \frac{-L}{E} l_\alpha \mathbf{r}_\alpha + \frac{-N}{G} l_\beta \mathbf{r}_\beta + I_n\mathbf{n}$$

by substituting from 2.19 and 2.20.

Hence

$$\nabla_3 I \cdot \mathbf{v} = -L l_\alpha v_\alpha - N l_\beta v_\beta + I_n v_n$$

Considering the other term:

$$(\nabla \times \mathbf{v}) \times \mathbf{n}$$

$$= \left[\frac{1}{A}\left(\frac{\partial}{\partial \alpha}(G v_\beta) - \frac{\partial}{\partial \beta}(E v_\alpha)\right)\mathbf{n} + \frac{1}{A}(N v_\beta)\mathbf{r}_\alpha - \frac{1}{A}(L v_\alpha)\mathbf{r}_\beta + (\nabla v_n)\times \mathbf{n}\right] \times \mathbf{n}$$

$$= \frac{1}{A^2}(N v_\beta)(-E)\mathbf{r}_\beta - \frac{1}{A^2}(L v_\alpha)(G)\mathbf{r}_\alpha + ((\nabla v_n)\cdot \mathbf{n})\mathbf{n} - (\nabla v_n)$$

$$= \frac{-N}{G} v_\beta \mathbf{r}_\beta + \frac{-L}{E} v_\alpha \mathbf{r}_\alpha - \frac{1}{E}\frac{\partial v_n}{\partial \alpha}\mathbf{r}_\alpha - \frac{1}{G}\frac{\partial v_n}{\partial \beta}\mathbf{r}_\beta$$

Hence

$$\mathbf{l}\cdot((\nabla \times \mathbf{v}) \times \mathbf{n}) = -L l_\alpha v_\alpha - N l_\beta v_\beta - l_\alpha \frac{\partial v_n}{\partial \alpha} - l_\beta \frac{\partial v_n}{\partial \beta}$$

Thus, on collecting the above expressions and cancelling the tangential velocity terms, 3.1 becomes –

The Intensity Rate Constraint :

$$\dot{I} + I_n v_n + l_\alpha \frac{\partial v_n}{\partial \alpha} + l_\beta \frac{\partial v_n}{\partial \beta} = 0 \qquad (3.2)$$

If the terms are considered at a point of maximal intensity, where the normal is in the direction of the light source, it is immediately apparent that l_α and l_β are both zero. I_n is also zero, as both I_α and I_β are zero and the gradient in the direction of the viewer must be zero. Consequently all the coefficients of the partial differential equation in v_n are zero, (as is the intensity rate at the image point), and locally it poses no constraint on v_n. The maximal intensity point was also significant in the shape from shading and stereo contour work of (Blake et al, 1985), where uniqueness (up to inversion) of the shape was proved under the condition that only one maximal point exists within the patch. The work in this thesis is done under the presumption that the shape could be found from such a technique, and subsequently will assume, when interpreting the intensity rate in terms of a three-dimensional velocity field, that the normal velocity can be found from the intensity rate constraint and the bounding contour velocity. Thus, for example, the case of a planar patch with the normal in the direction of the light source will be ignored.

The example surfaces from the Background Chapter, the circular cylinder and the torus, may be used to illustrate the Intensity Rate Constraint. The former will be considered first.

As the circular cylinder is identical in cross-section along the x-axis, it may be considered analogous to the circular curve of the Background Chapter which was considered in bending and stretching motion, Figures 2–2 and 2–3. In the latter case, the intensity rate constraint is rather trivially true, as the intensity rate is everywhere zero, (because the surface normal at any point on the circle, which deforms onto itself, is unchanged), as is the normal velocity. In the bending case, if it is considered viewed from the x-direction, then, reverting

to the planar curve notation, the intensity rate at a fixed image point may be derived as:

$$\dot{I} = -\frac{\dot{r}}{r} l_t \tan\theta$$

Given that the intensity gradient in the x-direction is zero, the gradient in the normal direction I_n may be derived from the gradient along the surface as:

$$I_n = \frac{\tan\theta}{r} l_t$$

Thus, the intensity rate constraint (simplified by omitting the terms in the α direction) evaluates to:

$$-\frac{\dot{r}}{r} l_t \tan\theta + \frac{\tan\theta}{r} l_t \dot{r}(1-\cos\theta) + l_t \frac{\partial}{\partial s}(\dot{r}(1-\cos\theta)) = 0$$

Now, two motions of the torus will be considered, the first the uniform expansion previously described, the second that of translating with unit speed in the x-direction, which may be parameterised as:

$$\mathbf{r}(\alpha,\beta,t) = ((a + r\cos\beta)\sin\alpha + t, r\sin\beta, (a + r\cos\beta)\sin\alpha) \qquad (3.3)$$

with corresponding velocities:

$$\begin{aligned} v_\alpha &= \frac{\cos\alpha}{a + r\cos\beta} \\ v_\beta &= \frac{-\sin\alpha\sin\beta}{r} \\ v_n &= \sin\alpha\cos\beta \end{aligned}$$

As both cases will be presumed to be viewed from the z-direction, the intensity gradient I_n is the same in both cases, so the derivation will be given as an example. The other two components of the intensity gradient are found thus:

$$I_\alpha = \frac{1}{E}\frac{\partial}{\partial\alpha}(\mathbf{l}\cdot\mathbf{n}) = -\frac{L}{E} l_\alpha$$

$$I_\beta = \frac{1}{G}\frac{\partial}{\partial\beta}(\mathbf{l}\cdot\mathbf{n}) = -\frac{N}{G} l_\beta$$

Letting $\mathbf{k}$ denote the unit vector in the z-direction, I_z (which is zero) is given by:

$$I_z = \nabla_3 I \cdot \mathbf{k}$$

$$= \left(-\frac{L}{E} l_\alpha \mathbf{r}_\alpha - \frac{N}{G} l_\beta \mathbf{r}_\beta + I_n \mathbf{n}\right)$$

$$\cdot \left(\frac{-\sin\alpha}{\sqrt{E}} \mathbf{r}_\alpha - \frac{\cos\alpha \sin\beta}{\sqrt{G}} \mathbf{r}_\beta + \cos\alpha\cos\beta \mathbf{n}\right)$$

$$= \frac{L l_\alpha}{\sqrt{E}} \sin\alpha + \frac{N l_\beta}{\sqrt{G}} \cos\alpha \sin\beta + I_n \cos\alpha\cos\beta$$

Hence

$$I_n = \frac{-L l_\alpha}{\sqrt{E}} \tan\alpha \sec\beta - \frac{N l_\beta}{\sqrt{G}} \tan\beta$$

$$= l_\alpha \tan\alpha + l_\beta \tan\beta$$

upon substituting for the first and second order magnitudes.

Now considering the uniform dilatation motion[1],the intensity rate at a fixed image point may be derived as:

$$\dot{I} = -\Theta\left((r + a\cos\beta) l_\alpha \tan\alpha + r l_\beta \tan\beta\right)$$

Substitution of the terms into the intensity rate constraint yields:

$$-\Theta\left((r + a\cos\beta) l_\alpha \tan\alpha + r l_\beta \tan\beta\right) + \left(l_\alpha \tan\alpha + l_\beta \tan\beta\right)\Theta(a\cos\beta + r)$$

$$+ l_\alpha \frac{\partial}{\partial\alpha}\left(\Theta(a\cos\beta + r)\right) + l_\beta \frac{\partial}{\partial\beta}\left(\Theta(a\cos\beta + r)\right)$$

Dividing by Θ and collecting terms in l_α and l_β gives:

$$\left(-(r + a\cos\beta)\tan\alpha + \tan\alpha(a\cos\beta + r)\right) l_\alpha$$

$$+ \left(-r\tan\beta + \tan\beta(a\cos\beta + r) - a\sin\beta\right) l_\beta$$

which can be seen to be zero as required.

The case of the translating torus has a particularly straightforward intensity rate to derive, as (similarly to the expanding torus) the normal does not change

[1]The motion considered here is that defined in (2.31).

as a function of α and β, thus for a fixed image point, as the velocity in the x-direction is 1

$$\dot{\alpha} = -\frac{\partial \alpha}{\partial x} = -\frac{1}{\frac{\partial x}{\partial \alpha}}$$

$$= -\frac{1}{(a + r\cos\beta)\cos\alpha}$$

Thus, for the fixed image point

$$\dot{I} = \mathbf{l} \cdot \dot{\mathbf{n}} = \mathbf{l} \cdot \mathbf{n}_\alpha \dot{\alpha}$$

$$= -\frac{\cos\beta}{\cos\alpha} l_\alpha$$

The intensity rate constraint may now be expanded as:

$$-\frac{\cos\beta}{\cos\alpha} l_\alpha + (l_\alpha \tan\alpha + l_\beta \tan\beta)\sin\alpha\cos\beta + l_\alpha \frac{\partial}{\partial\alpha}(\sin\alpha\cos\beta) + l_\beta \frac{\partial}{\partial\beta}(\sin\alpha\cos\beta)$$

which sums to zero.

3.6 An Extra Dimension to the Aperture Problem

If the intensity rate constraint and the motion constraint equation, used in the calculation of image flow, are compared, it is immediately apparent that they are closely related (as should be expected). In the forms 3.1 and 2.33 it is apparent that the former is a scaled up version of the latter, with an extra term expressing the change of intensity at a surface point, which is of course zero in the calculation of image flow. However, whereas the image flow "aperture problem" (*i.e.* the local indeterminacy of image flow in a direction of zero image intensity gradient) is most apparent in the compact form 2.33, the surface motion "aperture problem", the lack of any local information about either component of tangential velocity is most apparent from the expanded form 3.2. (Note that whereas the motion constraint equation does completely locally determine the normal component of velocity, the intensity rate constraint, being

a partial differential equation, does not). This new form of the aperture problem is a more general version of the oft-quoted example of viewing a constant albedo Lambertian sphere rotating about an axis through its centre (Horn & Schunck, 1981) – an object which is in motion but causes no change in the image.

It should be intuitively clear that this is a general problem in the case of a surface with reflectivity which is purely a function of $(\mathbf{l} \cdot \mathbf{n})$, and is not restricted to Lambertian reflectivity, orthographic projection, or an infinitely distant point light source, as used here. It follows straightforwardly from the fact that tangential motion moves a surface element onto the previous position of a surface element to which it has equivalent optical properties, which are independent of any rotation of the element about the surface normal. (Note that were the reflectivity to depend on the element's orientation about the surface normal, hence not be purely a function of $(\mathbf{l} \cdot \mathbf{n})$, the perceived intensity could be affected by tangential motion, even although the whole surface had the same reflectivity function.)

The more general form of the problem may be expressed mathematically as follows:

Consider a perceived intensity function $I(\mathbf{l} \cdot \mathbf{n}, \mathbf{r})$ depending upon the magnitude and direction relative to the surface normal of the illumination, and the position, the latter dependence encoding the effects of the projection, *e.g.* the decrease in intensity with $\cos^4$ of the angle from the centre line of projection in the case of perspective projection (Horn, 1986). (The latter dependence also covers the somewhat bizarre case of a more general dependence of reflectivity upon position, but perhaps such possibilities as idealised chameleons are too esoteric to ponder further!) Let the lighting be static, the magnitude in any direction depending only upon position. Consider the perceived intensity (and consequent gradient and rate) due to the illumination from a single direction, $\mathbf{l}$:

$$I = I(\mathbf{l} \cdot \mathbf{n}, \mathbf{r})$$

Hence:

$$\frac{dI}{dt} = \frac{\partial I}{\partial(\mathbf{l}\cdot\mathbf{n})}\left(\frac{d\mathbf{l}}{dt}\cdot\mathbf{n} + \mathbf{l}\cdot\frac{d\mathbf{n}}{dt}\right) + \frac{\partial I}{\partial \mathbf{r}}\cdot\frac{d\mathbf{r}}{dt}$$

$$= \frac{\partial I}{\partial(\mathbf{l}\cdot\mathbf{n})}\left(\left(\frac{d\mathbf{l}}{d\mathbf{r}}\frac{d\mathbf{r}}{dt}\right)\cdot\mathbf{n} + \mathbf{l}\cdot\frac{d\mathbf{n}}{dt}\right) + \frac{\partial I}{\partial \mathbf{r}}\cdot\frac{d\mathbf{r}}{dt}$$

$$= \frac{\partial I}{\partial(\mathbf{l}\cdot\mathbf{n})}\left(\left(\frac{d\mathbf{l}}{d\mathbf{r}}\mathbf{v}\right)\cdot\mathbf{n} + \mathbf{l}\cdot\frac{d\mathbf{n}}{dt}\right) + \frac{\partial I}{\partial \mathbf{r}}\cdot\mathbf{v}$$

Moreover

$$\nabla_3 I = \frac{\partial I}{\partial(\mathbf{l}\cdot\mathbf{n})}\left(\frac{d\mathbf{l}}{d\mathbf{r}}\cdot\mathbf{n} + \mathbf{l}\cdot\frac{\partial\mathbf{n}}{\partial\mathbf{r}}\right) + \frac{\partial I}{\partial \mathbf{r}}$$

Hence

$$\dot{I} + (\nabla_3 I\cdot\mathbf{v}) - \frac{dI}{dt} = 0$$

becomes

$$\dot{I} + \left[\frac{\partial I}{\partial(\mathbf{l}\cdot\mathbf{n})}\left(\frac{d\mathbf{l}}{d\mathbf{r}}\cdot\mathbf{n} + \mathbf{l}\cdot\frac{\partial\mathbf{n}}{\partial\mathbf{r}}\right) + \frac{\partial I}{\partial\mathbf{r}}\right]\cdot\mathbf{v}$$

$$-\left[\frac{\partial I}{\partial(\mathbf{l}\cdot\mathbf{n})}\left(\left(\frac{d\mathbf{l}}{d\mathbf{r}}\mathbf{v}\right)\cdot\mathbf{n} + \mathbf{l}\cdot\frac{d\mathbf{n}}{dt}\right) + \frac{\partial I}{\partial\mathbf{r}}\cdot\mathbf{v}\right] = 0$$

simplifying to

$$\dot{I} + \frac{\partial I}{\partial(\mathbf{l}\cdot\mathbf{n})}\left(\left(\mathbf{l}\cdot\frac{\partial\mathbf{n}}{\partial\mathbf{r}}\right)\cdot\mathbf{v} - \mathbf{l}\cdot\frac{d\mathbf{n}}{dt}\right) = 0$$

A comparison of this expression and the terms in the intensity rate constraint shows that the above may be reduced to

$$\dot{I} + I_n v_n + \frac{\partial I}{\partial(\mathbf{l}\cdot\mathbf{n})}\left(l_\alpha\frac{\partial v_n}{\partial\alpha} + l_\beta\frac{\partial v_n}{\partial\beta}\right) = 0$$

This expression for the intensity rate due to the variable magnitude illumination in any one direction is independent of the tangential velocity and hence the result of integrating over all illumination directions will be an expression which is independent of the tangential velocities and this more general case will exhibit a form of "aperture problem". This generalisation of the intensity rate constraint will not be pursued further in this thesis, as the simpler case already contains the key points of interest – the information about the normal velocity and the lack thereof regarding the tangential velocity.

3.7 Conclusions

The intensity rate constraint which relates the velocity at a point on a deforming uniform albedo Lambertian surface to the image intensity rate has been derived and shown to contain a scaled-up form of the well-known "aperture problem", in that the intensity rate yields no information about the tangential surface velocities. This problem also exists in more general cases.

Chapter 4

Interpretation of Cylindrical Deformation

4.1 Introduction

This section presents a technique for the interpretation of the changing intensity image generated by a cylindrical Lambertian surface undergoing a (possibly) non-rigid motion, such that it may be regarded as a deforming curve in a plane. The method, based around the intensity rate constraint developed in the previous chapter, uses a regularisation-style approach to estimate the tangential and normal velocities along the surface, assuming knowledge of the initial surface structure and the motion of the endpoints.

4.2 Interpreting the Intensity Rate Constraint

In the previous chapter, it was shown how the intensity rate constraint may be derived to relate the local temporal image intensity derivative, to the normal velocity of the corresponding surface point in the world. In this and the following chapter, a method is suggested by which the intensity rate may be interpreted in terms of a plausible three-dimensional velocity field over the surface which could have caused the perceived changing intensity image.

It is presumed that the instantaneous surface shape and position are known, as is the surface albedo and the position and strength of the light source. The

former information is considered to be calculable from the latter by a combination of stereo and shape-from-shading, presuming that the surface patch over which the three-dimensional velocity is sought is bounded by a contour for which the position at any point may be found by stereo, and within which a shape-from-shading method may be applied; see (Blake et al, 1985) for details of restrictions upon the class of surfaces to which this method may be applied.

Further, it is presumed that the three-dimensional velocity is known on the bounding contour, having been found from a form of stereo-motion technique. As the three-dimensional motion of a deforming curve is not uniquely determined by the two measurable image normal velocities, due to a further form of "aperture problem", this will itself require the application of a further constraint to the problem.

Given the analogy between the intensity rate constraint and the motion constraint equation, it seems appropriate to draw on the methods based on the latter for interpreting the image intensity rate in terms of the image flow as a source of inspiration for methods by which the intensity rate constraint may be used to interpret the image intensity rate in terms of the surface velocity field. As the methods for interpreting the intensity rate in terms of image flow are deriving only one tangential component of velocity, a simplified case of surface shape and motion is considered for the interpretation of the intensity rate in terms of the surface velocity field. In this, the "cylindrical" case, the problem is reduced by a dimension and methods proposed for the interpretation of planar curve motion may be applied

4.3 The Cylindrical Case

If a cylindrical surface, with the generators aligned with the x axis, deforms so as to remain cylindrical, with the generators aligned with the x axis, then the motion may be considered to be that of a deforming curve in the $y - z$ plane. This situation is depicted schematically in Figure 4–1, and an illustration of

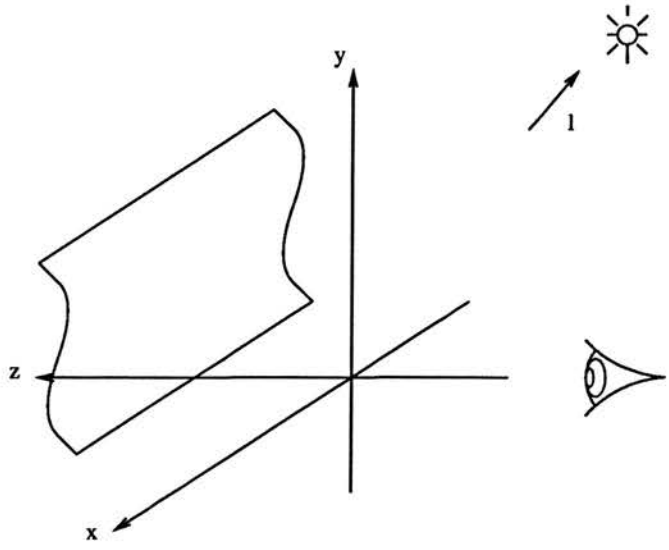

Figure 4–1: The cylindrical deformation case

the image produced by a circular cylinder thus aligned (illuminated from $\pi/4$ above the viewer) is shown in Figure 4–2.

As the motion has been reduced to that of a planar curve, and the $x-y$ plane is the plane which is most commonly considered, the subsequent examples will be presented as being motions in the $x-y$ plane, viewed along the x axis, as shown in Figure 4–3.

In the cylindrical case, parameterising the planar curve by its arc-length s, the intensity rate constraint may be simplified to:

$$\dot{I} + I_n v_n + l_t \frac{\partial v_n}{\partial s} = 0$$

where l_t is the tangential component of the light in the plane.

As in the case of contour motion in the image plane considered in the Background Chapter, the determination of the velocity field along the curve is ill-posed, and this leads to the suggestion that a form of regularisation should be applied, such as the velocity smoothness constraint of (Hildreth, 1984) or

Figure 4–2: Intensity pattern for circular cylinder appropriately viewed

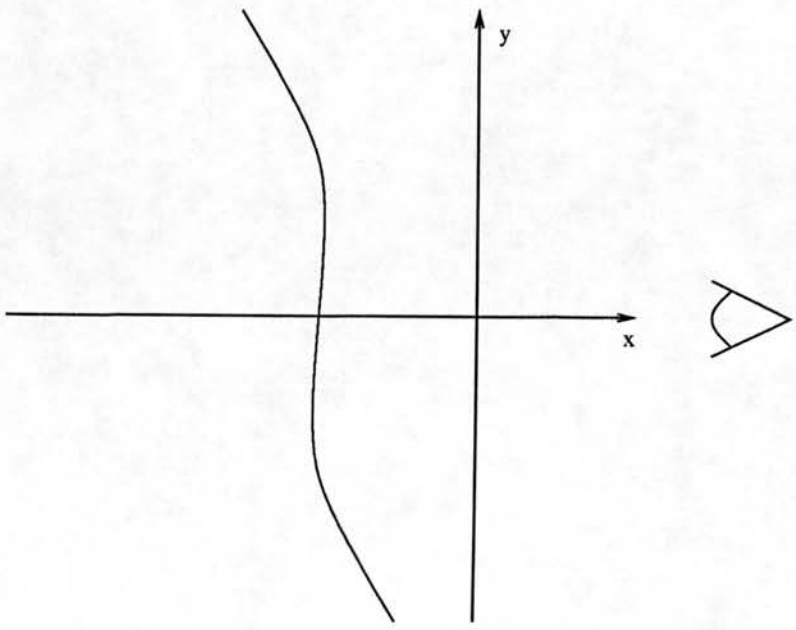

Figure 4–3: The cylindrical deformation case - in x-y plane

the stretch-minimising form of (D'Haeyer, 1986). The choice of regulariser will determine which forms of velocity field should be correctly found in the absence of noise. The velocity smoothness method will not correctly determine all forms of non-stretching motion, *i.e.* bending motions, whereas the stretch-minimising method will. As the stretching of the curve in this case is that in the physical world in which (perceptually) bending motion is a common and important form of motion, *e.g.* one perceives a piece of paper to bend without noticeably stretching, the method of stretch-minimising is used. (This is not intended to suggest that it is a preferable method for determining image curve motion, where the classes of motion which it is desirable to determine correctly are far less clear given the transformation produced by the projection, and the advantages of the velocity smoothness regulariser, such as not requiring a point at which the tangential velocity is known hence being applicable to closed contours, may predominate).

The stretch-minimising is done in the standard regularisation framework, v_t and v_n being determined to minimise the following, where λ is the regularisation parameter :

$$\int \left(\dot{I} + I_n v_n + l_t \frac{\partial v_n}{\partial s} \right)^2 + \lambda \left(\frac{\partial \mathbf{v}}{\partial s} \cdot \mathbf{t} \right)^2 ds$$

or re-expressing the regulariser:

$$\int \left(\dot{I} + I_n v_n + l_t \frac{\partial v_n}{\partial s} \right)^2 + \lambda \left(\frac{\partial v_t}{\partial s} - \kappa v_n \right)^2 ds \tag{4.1}$$

This expression can be minimised directly using numerical methods, as may be preferable in practice. Alternatively, formally minimising such an integral is a problem in the calculus of variations, (see *e.g.* (Stephenson, 1973)). Minimising $\int f(x, y, y')\, dx$, (where y' is the differential of y with respect to x) yields the Euler equation :

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'} \right) = 0$$

Applying this to 4.1, minimising with respect to v_t and v_n respectively gives

$$\frac{\partial}{\partial s} \lambda \left(\frac{\partial v_t}{\partial s} - \kappa v_n \right) = 0$$

and

$$I_n\left(\dot{I}+I_n v_n+l_t\frac{\partial v_n}{\partial s}\right)-\kappa\lambda\left(\frac{\partial v_t}{\partial s}-\kappa v_n\right)$$

$$-\frac{\partial}{\partial s}\left(l_t\left(\dot{I}+I_n v_n+l_t\frac{\partial v_n}{\partial s}\right)\right)=0$$

hence

$$\left(\frac{\partial^2 v_t}{\partial s^2}-\frac{\partial \kappa}{\partial s}v_n-\kappa\frac{\partial v_n}{\partial s}\right)=0$$

and

$$I_n\left(\dot{I}+I_n v_n+l_t\frac{\partial v_n}{\partial s}\right)-\kappa\lambda\left(\frac{\partial v_t}{\partial s}-\kappa v_n\right)$$

$$-\frac{\partial l_t}{\partial s}\left(\dot{I}+I_n v_n+l_t\frac{\partial v_n}{\partial s}\right)$$

$$-l_t\left(\frac{\partial \dot{I}}{\partial s}+\frac{\partial I_n}{\partial s}v_n+\left(I_n+\frac{\partial l_t}{\partial s}\right)\frac{\partial v_n}{\partial s}+l_t\frac{\partial^2 v_n}{\partial s^2}\right)=0$$

(Note that a term from the last equation was erroneously omitted in (Cameron-Jones, 1988)).

This is a pair of coupled second order differential equations in v_t and v_n, for which a numerical method of solution may be attempted, given the boundary values of v_t and v_n. The use of central finite difference approximations leads to a sparse system of linear equations, which would be best solved by a method exploiting the near-diagonal form; however a more general partial pivoting method (from (Conte & deBoor, 1981)) was used in the implementation to allow for the possible use of other types of constraint, which might change the structure of the equations. For further details regarding the issue of numerical implementation see Appendix A, which describes in more detail the implementation of the more general case (of the next chapter), of which this is a subset.

Before results from solving some cases numerically are given, an example bending motion will be used to illustrate that the analytic expression for the velocity field is a solution of the above pair of differential equations. This example motion will be the circular bending motion already met in the previous chapters.

Substituting for the terms in the first equation yields:

$$-\frac{\dot{r}}{r^2}\sin\theta - \frac{-1}{r}\frac{\dot{r}}{r}\sin\theta = 0$$

If the terms in the second equation are considered, it can be seen that those which contain the intensity rate constraint, ($I_n(...)$ and $-\frac{\partial l_t}{\partial s}(...)$), are zero, as it has already been shown to be satisfied by this motion in the previous chapter. The term involving λ has also already been shown to be zero (in the Background Chapter). Thus all that remains is to demonstrate that we have the correct expressions for the last term. Disregarding the factor $-l_t$, this is:

$$\frac{\dot{r}}{r^2}\left(l_n\tan\theta - l_t\sec^2\theta\right) + \frac{1}{r^2}\left(l_t\sec^2\theta - l_n\tan\theta\right)\dot{r}(1-\cos\theta)$$

$$+\left(\frac{1}{r}l_t\tan\theta - \frac{1}{r}l_n\right)\frac{\dot{r}}{r}\sin\theta + l_t\frac{\dot{r}}{r^2}\cos\theta$$

which sums to zero.

A pair of test cases, both based on a circular cylinder deforming so as to remain circular, using perfect input data, are illustrated in Figures 4–4 to 4–6. (The next chapter which tackles the full dimensional case will include results demonstrating the deterioration of the method with noise). In the first case, the cylinder bends without stretching, in the second it stretches uniformly along itself, hence the intensity rate and normal velocity are zero. In all these figures the correct velocities are shown on the left and the estimated velocities on the right. The bending case is illustrated twice, once (Figure 4–4) with the velocities superposed upon the intensity pattern seen by the viewer, and once (Figure 4–5) with the velocities shown in cross-section through the cylinder. The stretching case is only illustrated in the latter form.

The two forms of curve motion are firstly the bending motion:

$$\mathbf{r}(s,t) = \left((r+\dot{r}t)\left(\cos\left(\frac{s}{(r+\dot{r}t)}\right)-1\right),(r+\dot{r}t)\sin\left(\frac{s}{(r+\dot{r}t)}\right)\right)$$

and secondly the stretching motion:

$$\mathbf{r}(s,t) = \left(r\left(\cos\left(\frac{s+et}{r}\right)-1\right),r\sin\left(\frac{s+et}{r}\right)\right)$$

both considered only at the moment where $t = 0$. (In the second case, s is the arc-length only at $t = 0$). $\dot{r}$ and e are the rate of radius increase and the rate of extension respectively.

Two instances of both these motions were tested with a range of values of λ. Letting θ denote the angle $\frac{s}{r}$ and $\Delta\theta$ the increment used, the cases tried were:

Case	l_x	l_y	ρ	r	$\dot{r}$ or e	$\theta \geq$	$\theta \leq$	$\Delta\theta$
Bending A	0.7	0.7	255	400	400	-0.5	0.5	0.1
Bending B	0.7	-0.7	255	400	-800	-0.25	0.25	0.05
Stretching A	0.7	0.7	255	400	0.2	-0.5	0.5	0.1
Stretching A	0.7	-0.7	255	800	-0.1	-0.75	0.75	0.15

Table 4–1: Cylindrical Motion Cases – Parameter Values

The two A cases with $\lambda = 0$ were used to produce Figures 4–4 to 4–6. (Note that the boundary velocities are not shown in the figures as they are inputs to the method not outputs from it).

A table demonstrating the result of varying the regularisation parameter in the four cases is given below. The velocity errors are given as the root mean square (rms) over the points at which the velocity is calculated. The error at an individual point is the magnitude of the vector between the (known) true velocity and the estimated velocity at the point. The rms of the true velocities over the same sample points is given for comparison.

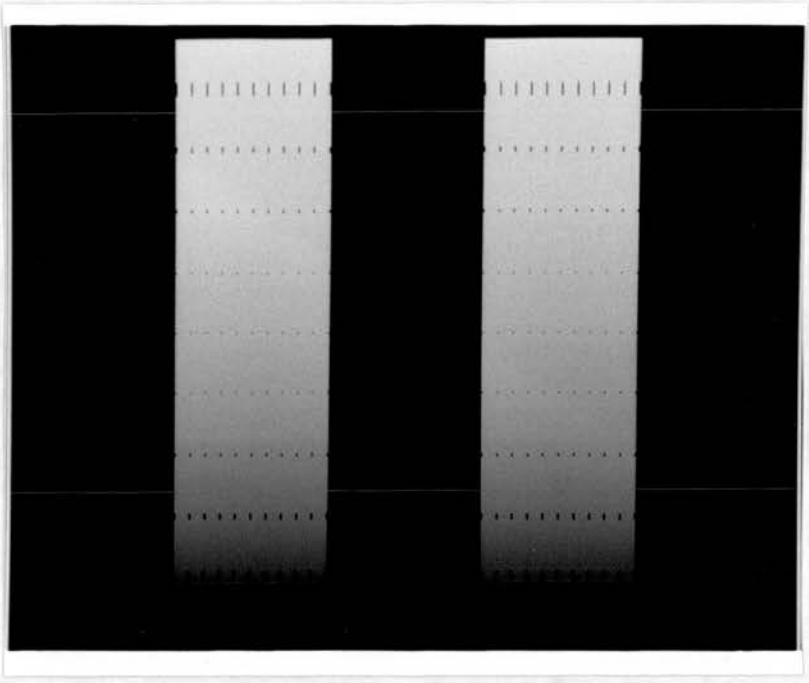

Figure 4–4: Cylindrical bending motion – velocities superposed on image ($\lambda = 1$)

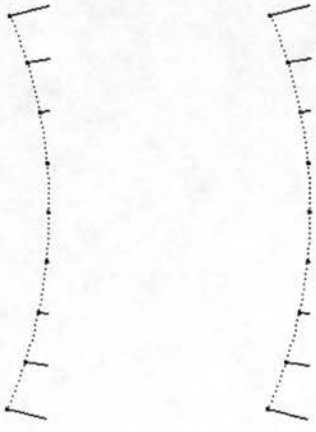

Figure 4–5: Circular bending velocities - true and estimated ($\lambda = 1$)

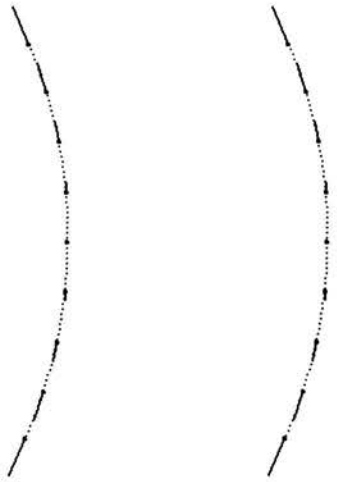

Figure 4–6: Circular stretching velocities - true and estimated ($\lambda = 1$)

Case	Bending A	Bending B	Stretching A	Stretching B
True Velocity (rms)	1.7671e+01	8.8609e+00	2.0656e+01	3.0984e+01
$\log_{10}\lambda$	Velocity Errors(rms)			
10	7.9475e-01	3.8121e-01	9.0736e+01	9.0962e+01
8	7.8579e-01	3.7413e-01	9.0457e+01	9.0861e+01
6	5.8777e-01	1.5668e-01	7.2278e+01	8.4537e+01
4	1.3856e-02	1.3872e-03	3.4330e+00	1.2211e+01
2	3.0917e-02	3.9525e-03	3.5666e-02	1.4110e-01
0	3.1186e-02	3.9795e-03	3.5680e-04	1.4132e-03
-2	3.1188e-02	3.9797e-03	3.5680e-06	1.4132e-05
-4	3.1188e-02	3.9798e-03	3.5680e-08	1.4132e-07
-6	3.1188e-02	3.9798e-03	3.5680e-10	1.4132e-09
-8	3.1188e-02	3.9798e-03	3.5676e-12	1.4132e-11
-10	3.1188e-02	3.9798e-03	3.6252e-14	1.4213e-13

The examples of bending motion illustrate that, as the correct value of the regularising term is zero, the solution found best matches the true veloc-

ity when λ is such that the discretisation error in neither term predominates. The examples of uniform stretching show that the solution found tends to the true velocity as the regularising parameter is decreased, causing the method to become similar to minimising the stretch term subject to the intensity rate equation as a constraint, when the uniform stretch case should be solved correctly. The great accuracy achieved is not to be expected in a general case of stretching motion, but reflects the fact that in this case the true (zero) normal velocity is found from the intensity rate constraint and the tangential velocity may be found exactly from the discretisation of the partial differential equation arising from the minimisation of the stretch term subject to the constraint of zero normal velocity.

4.4 Conclusion

It has been shown how, in the reduced dimensional case of (suitably viewed) cylindrical motion of a surface, the intensity rate constraint derived in the previous chapter may be combined with an assumption of stretch minimisation in a regularising framework to enable the estimation of surface velocities from the changing intensity image. Given perfect input data, these estimates match the true velocities in such important cases as pure bending motion, and uniform stretching motion.

Chapter 5

Interpretation of Smooth Surface Deformation

5.1 Introduction

In this chapter the method of the previous chapter for interpreting the image intensity change of a cylindrical deformation is extended to the case of a (potentially) doubly-curved constant albedo Lambertian surface undergoing a smooth deformation. A (square of) divergence term is used as a regulariser to infer the velocity field over the surface, assuming a knowledge of the initial surface structure and the velocity on the bounding contour. The method is tested on ideal synthetic data as a check on the analysis and on noisy synthetic data to demonstrate the (presently unattainable) typical level of accuracy to which the intensity rate would have to be measured for practical application of the method.

5.2 The Divergence-Based Regulariser

It was shown in Chapter 3 that the intensity rate at a point in an image of a deforming constant albedo Lambertian surface is related to the normal velocity and its derivatives by the intensity rate constraint 3.2, and it was proposed in Chapter 4 that in a special reduced dimensional case of surface motion, equivalent to curve motion in the plane, the velocity field over the

surface may be estimated by the application of a stretch-minimising regulariser. When considering the generalisation of the method of the previous chapter to the unrestricted form of smooth surface deformation, the most straightforward approach is to choose a regularisation term which is again a measure of surface stretching, and apply it to the full form of the intensity rate constraint.

The form of this regulariser should be such that it is zero in the case of a pure bending motion and hence the regularisation will yield the correct solution given correct input data representing this significant case. In a pure bending motion the surface dilatation and shear are both zero, hence terms representing either or both seem plausible candidates. The dilatation (per unit area) is measured by the divergence, a differential invariant, which has the prerequisite property of depending upon both tangential and normal velocities and (as will be shown below) is mathematically convenient for considering the limiting case of ideal input data where λ may be made very small. Thus the (square of) divergence was used as the regulariser in this work; however, if a similarly appropriate shear-based term were found it might yield correct results in some other interesting cases.

Thus, the velocity field is chosen by minimising (with respect to $\mathbf{v}$) the following integral over the surface, (subject to the known $\mathbf{v}$ on the bounding contour):

$$\int\int A\left(\dot{I} + I_n v_n + l_\alpha \frac{\partial v_n}{\partial \alpha} + l_\beta \frac{\partial v_n}{\partial \beta}\right)^2 + A\lambda\,(\nabla \cdot \mathbf{v})^2\, d\alpha d\beta$$

As commented above, this method should yield the correct velocity field in the case of a pure bending motion of a surface (which is completely visible and unshadowed), independent of the magnitude of λ. A further significant case is that in which the surface is undergoing a uniform expansion (with everywhere constant divergence), and λ is sufficiently small that the regularisation results in minimising the (square of) divergence term over the surface, subject to the normal velocity found from the intensity rate constraint and the known velocity on the bounding contour. In this case the "divergence theorem" (page 239 of

(Weatherburn, 1931)) may be applied to the known velocities to show that the minimisation is consistent with the correct motion being found.

The divergence theorem states that for a closed curve on a surface (arclength s), letting $\mathbf{m}$ be the unit vector tangential to the surface and normal to the curve in the direction out of the enclosed region, the surface integral of the divergence of a vector quantity, such as $\mathbf{v}$ is given thus:

$$\int\int \nabla \cdot \mathbf{v} dS = \int \mathbf{v} \cdot \mathbf{m} ds - \int\int J\mathbf{v} \cdot \mathbf{n} dS \tag{5.1}$$

Hence the integral of the divergence over the surface may be found from the normal velocity (which is found from the intensity rate constraint) and the known velocity on the bounding contour, which are constraints upon the minimisation of the integral of the squared divergence. Consequently the minimisation of the integral of the square of the divergence is done subject to this (implicit) constraint and hence the result of that minimisation will be the correct uniform value of divergence and thus the correct motion field.

As in the previous chapter, the minimisation of the regularised problem is performed by applying the calculus of variations to the integral, which may first be re-expressed by expanding the divergence term in terms of the velocity components:

$$\int\int A\left(\dot{I} + I_n v_n + l_\alpha \frac{\partial v_n}{\partial \alpha} + l_\beta \frac{\partial v_n}{\partial \beta}\right)^2$$

$$+ A\lambda\left(\frac{A_\alpha}{A} v_\alpha + \frac{\partial v_\alpha}{\partial \alpha} + \frac{A_\beta}{A} v_\beta + \frac{\partial v_\beta}{\partial \beta} - J v_n\right)^2 d\alpha d\beta \tag{5.2}$$

Applying the calculus of variations yields three linked second order partial differential equations of which two are the result of differentiating the divergence term by the surface coordinates:

$$\frac{A_{\alpha\alpha}A - {A_\alpha}^2}{A^2} v_\alpha + \frac{A_\alpha}{A}\frac{\partial v_\alpha}{\partial \alpha} + \frac{\partial^2 v_\alpha}{\partial \alpha^2} +$$

$$\frac{A_{\alpha\beta}A - A_\alpha A_\beta}{A^2} v_\beta + \frac{A_\beta}{A}\frac{\partial v_\beta}{\partial \alpha} + \frac{\partial^2 v_\beta}{\partial \alpha \partial \beta} - J_\alpha v_n - J\frac{\partial v_n}{\partial \alpha} = 0$$

$$\frac{A_{\alpha\beta}A - A_\alpha A_\beta}{A^2} v_\alpha + \frac{A_\alpha}{A}\frac{\partial v_\alpha}{\partial \beta} + \frac{\partial^2 v_\alpha}{\partial \alpha \partial \beta} +$$

$$\frac{A_{\beta\beta}A - {A_\beta}^2}{A^2}v_\beta + \frac{A_\beta}{A}\frac{\partial v_\beta}{\partial \beta} + \frac{\partial^2 v_\beta}{\partial \beta^2} - J_\beta v_n - J\frac{\partial v_n}{\partial \beta} = 0$$

The third equation has more terms:

$$AI_n\left(\dot{I} + \ldots\right) - \lambda AJ\left(\nabla \cdot \mathbf{v}\right)$$

$$-\frac{\partial}{\partial \alpha}\left(Al_\alpha\left(\dot{I} + \ldots\right)\right) - \frac{\partial}{\partial \beta}\left(Al_\beta\left(\dot{I} + \ldots\right)\right) = 0 \qquad (5.3)$$

The resulting expression may be simplified by observing that as the vector **l** is fixed, its divergence is zero hence:

$$\left(\frac{\partial}{\partial \alpha}\left(Al_\alpha\right) + \frac{\partial}{\partial \beta}\left(Al_\beta\right)\right) = AJl_n$$

Hence expanding 5.3, dividing by A, and collecting terms yields:

$${l_\alpha}^2\frac{\partial^2 v_n}{\partial \alpha^2} + 2l_\alpha l_\beta\frac{\partial^2 v_n}{\partial \alpha \partial \beta} + {l_\beta}^2\frac{\partial^2 v_n}{\partial \beta^2} + \left(Jl_n l_\alpha + l_\alpha\frac{\partial l_\alpha}{\partial \alpha} + l_\beta\frac{\partial l_\alpha}{\partial \beta}\right)\frac{\partial v_n}{\partial \alpha}$$

$$+\left(Jl_n l_\beta + l_\alpha\frac{\partial l_\beta}{\partial \alpha} + l_\beta\frac{\partial l_\beta}{\partial \beta}\right)\frac{\partial v_n}{\partial \beta} + \left(\left(Jl_n - I_n\right)I_n - \lambda J^2 + l_\alpha\frac{\partial I_n}{\partial \alpha} + l_\beta\frac{\partial I_n}{\partial \beta}\right)v_n$$

$$+\lambda J\frac{\partial v_\alpha}{\partial \alpha} + \lambda J\frac{A_\alpha}{A}v_\alpha + \lambda J\frac{\partial v_\beta}{\partial \beta} + \lambda J\frac{A_\beta}{A}v_\beta + \left(Jl_n - I_n\right)\dot{I} - l_\alpha\frac{\partial \dot{I}}{\partial \alpha} - l_\beta\frac{\partial \dot{I}}{\partial \beta}$$

$$= 0 \qquad (5.4)$$

These three coupled second order partial differential equations may be solved numerically over a bounded surface region, given the velocity on the bounding contour. Some example results from the application of the method to synthesised data are given in the next section. Aspects of the numerical implementation used are outlined in Appendix A.

The manual substitution of analytic expressions for example velocity fields, with associated intensity rates and intensity gradients, into these PDEs to demonstrate that, in a given case, the veridical velocity field is a solution of the PDEs seems impractical in cases of significant interest. The expressions used in the implementation programs were checked in a limited sense by ensuring that over a range of parameters the numerical sum of the equations in unstretching cases was small (*e.g.* of order 10^{-8}) relative to the largest term

(*e.g.* $l_\alpha{}^2 \frac{\partial^2 v_n}{\partial \alpha^2}$). However, two cases will be considered at a coarse level. The first is the circular cylindrical bending used in previous chapters, the second a unit speed translation of a torus in the $x-$direction as also previously discussed.

In the cylindrical bending case, there is no variation of velocity or first or second order magnitudes with α; thus it is obvious that the first PDE is satisfied by the veridical solution. Consideration of the latter two PDEs in this case shows that they may be reduced to those used in the previous chapter for exclusively cylindrical motions (by substituting $J = \kappa$ *etc.*). Thus as the bending circular cylinder has been shown to satisfy the PDEs for the exclusively cylindrical case, it satisfies these more general ones.

In the case of the unit speed translating torus, it has already been shown that the intensity rate constraint is satisfied, and the divergence is clearly zero. From the latter, it follows that the first two PDEs will be satisfied as they are differentials of the divergence with respect to α and β. (The cancellation of the analytic expressions used in the programs was checked manually for these equations in this case). If the third PDE is considered in its pre-expanded form 5.3, it is clear that knowing the intensity rate constraint holds, as has been shown, and that the divergence is zero, leads to this being satisfied.

5.3 Application Results

The method was tested on data representing two deforming surfaces: (1) a bending circular cylindrical region (as used in the previous chapter in which the deformation was restricted to being cylindrical, and hence one dimension of the problem ignored) and (2) an expanding, rotating and translating toroidal region. For simplicity, results are presented for only one deformation of each of these surfaces with the following variations of the input:

1. Ideal input data including all spatial derivatives required, with temporal derivatives given exactly by the instantaneous rates of change.

2. As 1, except that the spatial derivatives of the intensity rate were found by finite difference approximation (as given in Appendix A).

3. As 2, except that Gaussian noise was added to the intensity rate.

4. As 2, except that Gaussian noise was added to the input velocity components.

5. As 2, except that Gaussian noise was added to the presumed light source vector.

6. As 2, except that Gaussian noise was added to the instantaneous shape information.

7. As 2, except that the intensity rate and velocities were all determined by (forward) finite differencing over a finite time period.

8. As 5, except that Gaussian noise was added to the intensity rate.

As these tests were the first to investigate the possibility of extracting velocity information in this manner, there was no clear standard to draw upon for appropriate magnitudes of input noise. Consequently the experiments used a wide range of input magnitudes to establish only the gross order of magnitude of input accuracy required, rather than investigating the matter in fine detail. More complex modelling of camera noise as suggested by the experimental results of camera tests in (McIvor, 1990) seems unnecessary at this stage.

Before giving the results of these experiments the two cases of region deformation will be explained more fully.

The first, that of the circular bending motion is a variant of one used in the previous chapter, here illustrated in Figure 5–1. The surface, as a function of the surface coordinates and time, is given by the expression:

$$\mathbf{r}(\alpha,\beta,t) = \left(\alpha, (r+\dot{r}t)\sin\left(\frac{\beta}{(r+\dot{r}t)}\right), (r+\dot{r}t)\left(\cos\left(\frac{\beta}{(r+\dot{r}t)}\right)-1\right)\right)$$

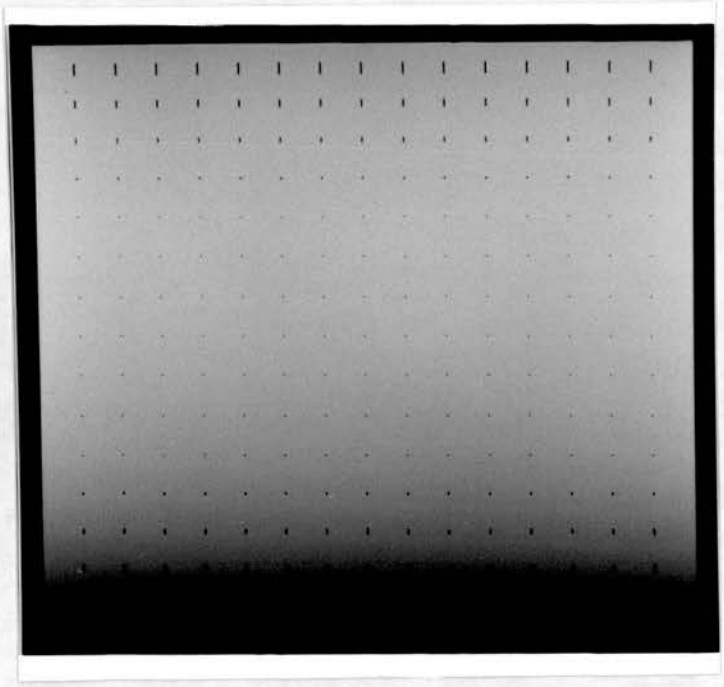

Figure 5–1: Cylindrical region bending - velocity field superposed on intensity image

In the example used, the surface region with $r = 300$, $\dot{r} = 80$, α and β in the range $[-240, 240]$, (both sampled every 30 except in cases where otherwise indicated) was viewed from the z direction, at time $t = 0$. (All dimensions are in pixels. In the case of orthographic projection, this appears a natural unit). The albedo was 255, and the light source vector at $(0.3, 0.3, 0.95)$.

The second case, that of the toroidal region expansion, rotation (about the z axis) and translation, (viewed from the y direction), is illustrated in Figures 5-2 to 5-5 where the independent cases of expansion, rotation and translation of the magnitudes used are shown, followed by the resulting velocity field.

The surface as a function of the surface coordinates and time is given (at time $t = 0$) by:

$$\mathbf{r}(\alpha, \beta, t) = ((1 + \Theta t)(a + r\cos\beta)\sin\alpha\cos\omega t - (1 + \Theta t)r\sin\beta\sin\omega t + v_x t,$$

$$(1 + \Theta t)r\sin\beta\cos\omega t + (1 + \Theta t)(a + r\cos\beta)\sin\alpha\sin\omega t,$$

$$(1 + \Theta t)(a + r\cos\beta)\cos\alpha)$$

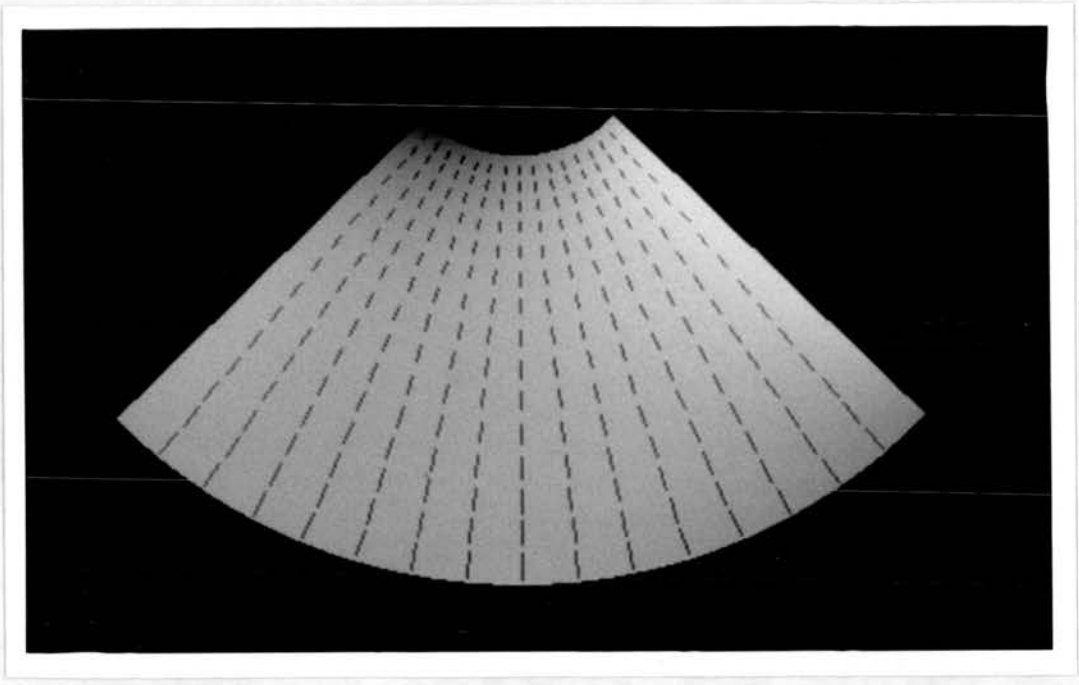

Figure 5–2: Toroidal region expansion - velocity field superposed on intensity image

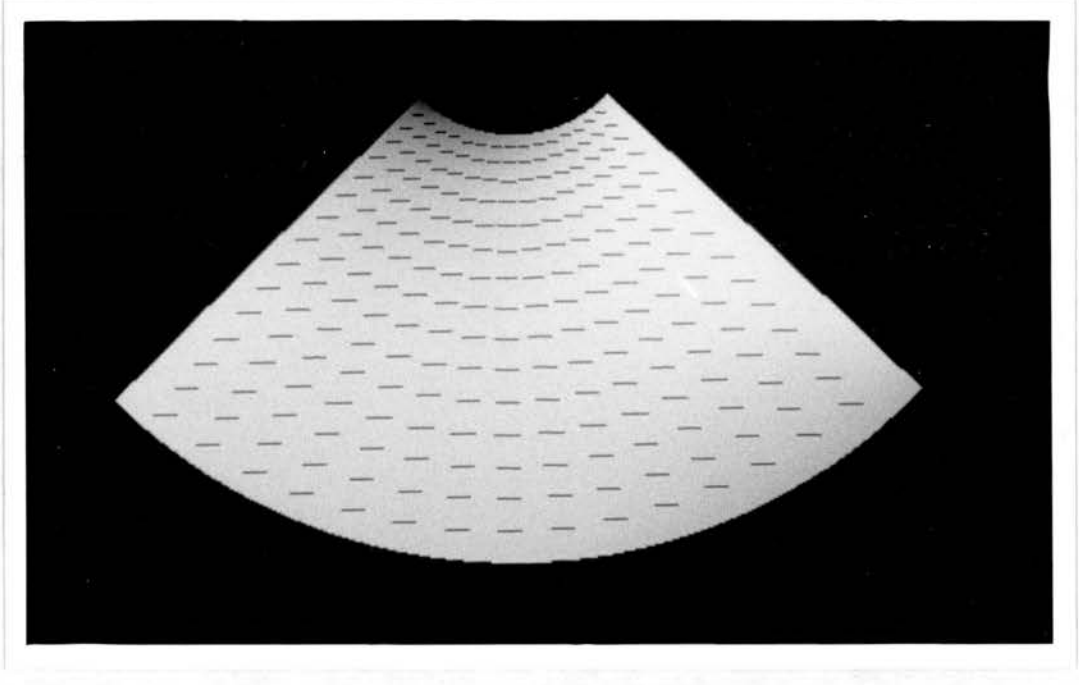

Figure 5–3: Toroidal region rotation - velocity field superposed on intensity image

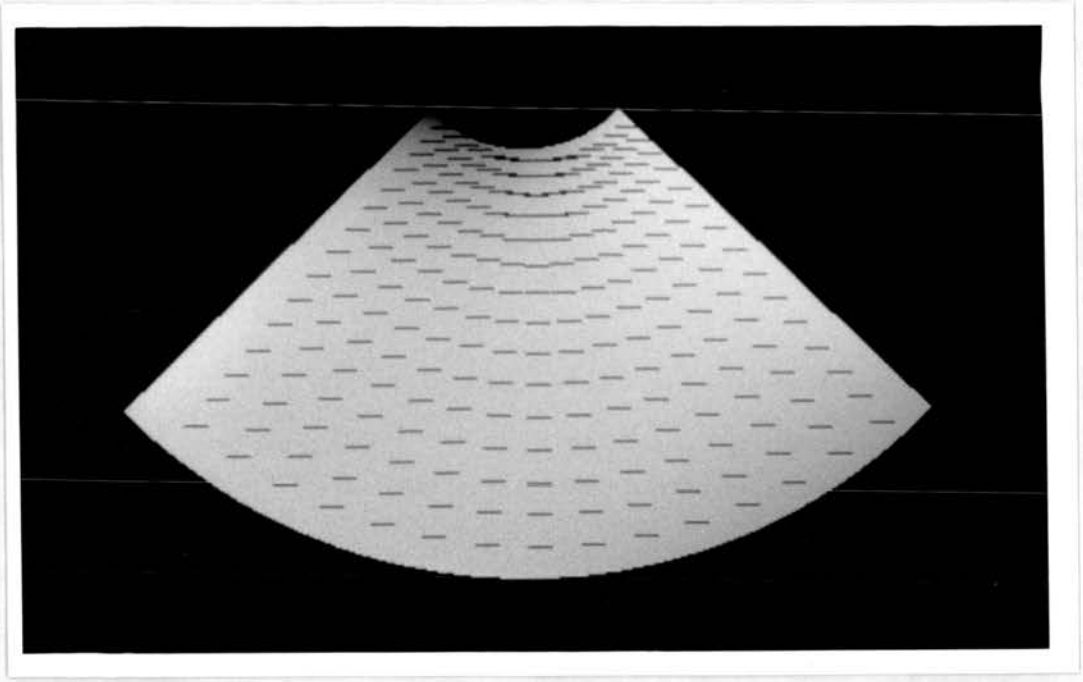

Figure 5–4: Toroidal region translation - velocity field superposed on intensity image

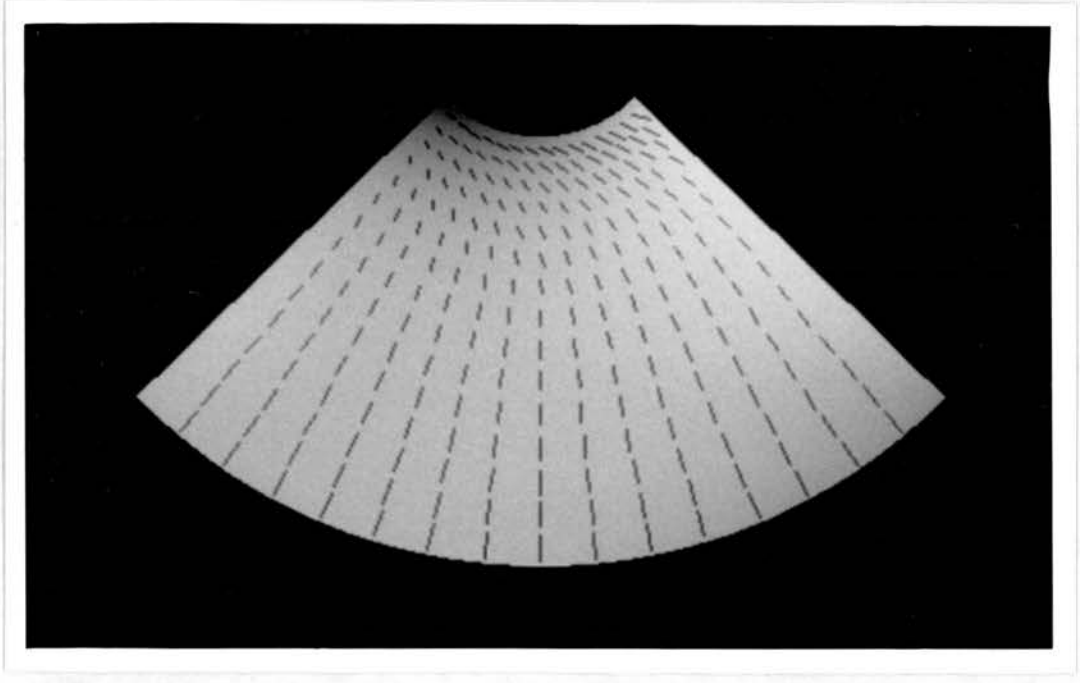

Figure 5–5: Toroidal region expansion, rotation and translation - velocity field superposed on intensity image

In the example used, the surface region with $a = 200$, $r = 150$, $\Theta = 0.05$, $\omega = 0.07$, $v_x = 10$, α in the range $[-0.8, 0.8]$, and β in the range $[1.2, 2.8]$ (both sampled every 0.1, except in cases where otherwise indicated) was viewed from the y direction, at time $t = 0$. The albedo was 255 and the light source vector at $(-0.3, 0.9, -0.3)$.

The two types of motion were both chosen to be cases in which the intensity at the (changing) image point corresponding to a moving point of the viewed surface changes; thus the standard assumption (see Chapter 2) that image flow is equivalent to optic flow does not hold. In the toroidal example, if the rotation is set to zero, the assumption does hold (and the method proposed in this chapter still works).

Further, the experiments were designed such that the regions include (between them) sampled points of positive, negative and zero Gaussian curvatures, and positive and negative mean curvatures. (The points of zero mean curvature on the toroidal patch are not sampled). The range of surface normals in the regions is such that the light source has been positioned close to the viewing direction in each case to ensure that the region is visible in the photographed intensity images (which seems to require a sharper contrast across the edge of the region than when viewing on a graphics terminal). Although this positioning of the light source might seem suspiciously convenient to those familiar with techniques which rely upon the light source being in the viewing direction, it should be made clear that the method proposed in this chapter is not dependent on such an alignment. Indeed, the magnitude of the main input signal, the intensity rate, is greater for a given rate of change of the angle between the surface normal and the light source when the light source is nearly tangential to the surface than nearly perpendicular.

The magnitudes of the motions concerned were chosen such that the velocities to be determined have a root mean square magnitude of about 10, which given the distance between the points in the images, enables the velocity vectors to be reasonably distinguished when displayed. (This is also true of the magnitudes of the individual expansion, rotation and translation motions which

combine to make up the overall toroidal motion). Hereafter, the overall toroidal motion will generally be referred to as an expansion to distinguish it from, for example, a rigid motion. The translation and rotation will be assumed.

The results which will be given are in the form of root mean square (rms) velocity errors, (calculated over the points at which the velocity is determined), which may be compared with the rms true velocity over the same set of points. The error at each individual point is the magnitude of the vector which is the difference between the (known) true and estimated three-dimensional velocity. In all cases where noise was added, the example was repeated 100 times and the resulting mean and standard deviation of the rms velocity errors will be given.

5.3.1 Application to Ideal Input Data

This section contains the results of applying the regularisation method to perfect input data representing the cases considered, *i.e.* the value of every coefficient of the partial differential equations is assumed to be exactly known at each point. This demonstrates the maximum accuracy which this numerical implementation of the method can achieve on these cases, as the error is entirely due to the discretisation of the partial differential equations. Consequently, it is against the results obtained in this section that those of the subsequent sections should be judged.

The results of applying the method to the cylindrical bending case and the toroidal stretching case with a wide range of regularisation parameters and a range of grid sizes from 4×4 increments to 48×48 increments are given in the subsequent tables. These grids have equal α and β increments ranging from 120 (cylindrical case) and 0.4 (toroidal case) down to 10 and 0.0333333. The cases in which the results are denoted by an asterisk are those in which the linear equations representing the discretisation of the partial differential equations were not able to be solved using pivoting with the compact representation of

the matrix implemented to save storage space — see Appendix A, section A.3, for further details.

Increment	120	60	30	15	10
True Velocity	5.20239	8.3905	9.87864	10.5991	10.8358
$\log_{10}\lambda$	Velocity Errors (rms)				
10	2.58600	0.715708	0.270714	*	*
8	2.58341	0.714172	0.261632	*	*
6	2.36858	0.639108	0.174729	*	*
4	0.200167	0.0509551	0.0148264	0.00352104	0.00184416
2	0.341902	0.0799144	0.0198046	0.00503911	0.00240085
0	0.345603	0.0806519	0.0199603	0.00502905	0.00238635
-2	0.345641	0.0806566	0.0199516	0.00501943	0.00237710
-4	0.345642	0.0806562	0.0199517	0.00501969	0.00237477
-6	0.345642	0.0806562	0.0199517	0.00501967	0.00237477
-8	0.345642	0.0806562	0.0199517	0.00501972	0.00237477
-10	0.345642	0.0806562	0.0199517	0.00501964	0.00237536

Table 5–1: Cylindrical Bending Case – Ideal Input Data

As can be seen, the bending case, in which the quantity which is being minimised is zero at the correct solution, has a more pronounced improvement in accuracy with the change in grid size, (the error being proportional to the square of the increment), and demonstrates a significant peak in accuracy with varying λ, similarly to the case in the previous chapter. Again in the bending case, the correct result should (in the continuous case) be found for any positive λ so the trade-off in discretisation error between the two terms in the regularisation should be the predominant factor in any variation of accuracy with λ. The stretching case can be seen to deteriorate at the finest grid size, illustrating a practical weakness of the simple numerical approach when using this form of regulariser – see Appendix A for further details of the approach. Similarly the less pronounced peak in accuracy with the variation of λ is expected because, in the stretching case, the correct result is found only as λ tends to zero. In

Increment	0.4	0.2	0.1	0.05	0.0333333
True Velocity	10.5283	10.944	11.1511	11.2544	11.2885
$\log_{10}\lambda$	Velocity Errors (rms)				
10	33.4799	90.8248	26.3972	*	*
8	34.2417	109.862	21.9868	*	*
6	17.0536	21.5462	24.5621	*	*
4	2.70798	2.56359	2.47420	*	*
2	0.242526	0.0804116	0.0583997	*	*
0	0.248366	0.0790501	0.0542144	0.0508469	0.0556922
-2	0.248443	0.0790897	0.054231	0.0504993	0.0571166
-4	0.248444	0.0790908	0.0542375	0.0506068	0.0571867
-6	0.248444	0.0790910	0.0542375	0.0506129	0.0564594
-8	0.248444	0.0790910	0.0542375	0.0506129	0.0564594
-10	0.248444	0.0790910	0.0542375	0.0506129	0.0564594

Table 5–2: Toroidal Stretching Case – Ideal Input Data

practice, this is the case where the point normal velocities are effectively found independently of the tangential velocities, because the coefficients of the point normal velocities in the linear equations that represent the discretisations of 5.4 are sufficiently greater than those of the point tangential velocities.

5.3.2 Intensity Rate Spatial Derivatives Found By Differencing

This section contains the results of applying the method to data which was identical to the previous section except that the intensity rate spatial derivatives were found by the application of finite difference approximations over a 3×3 neighbourhood of the intensity rate. This introduced the effect of spatial sampling of what may be considered the primary input of the method. (The other input which might be thus considered is the velocity, but the effect of spatial sampling was already present even in the "ideal" input). Similarly to

the previous section, these results form a benchmark against which those of the later sections may be judged. In order to facilitate the comparison between the results of this section and those of the previous section with the analytically determined spatial intensity derivatives, tables of the relative rms velocity errors are also given. The relative rms errors are the differences between the rms velocity errors in the corresponding cases, divided by the rms velocity error in the analytically determined spatial intensity derivative case.

Increment	120	60	30	15	10
True Velocity	5.20239	8.3905	9.87864	10.5991	10.8358
$\log_{10}\lambda$	Velocity Errors (rms)				
10	2.58599	0.715711	0.270726	*	*
8	2.58268	0.714509	0.262633	*	*
6	2.31996	0.646049	0.185626	*	*
4	0.768408	0.257404	0.0646034	0.0163933	0.00768538
2	1.21213	0.331827	0.0805598	0.0204464	0.00921298
0	1.21766	0.332754	0.0807580	0.0204535	0.00920586
-2	1.21772	0.332761	0.0807503	0.0204414	0.00918633
-4	1.21772	0.332761	0.0807503	0.0204412	0.00918505
-6	1.21772	0.332761	0.0807503	0.0204412	0.00918517
-8	1.21772	0.332761	0.0807503	0.0204412	0.00918446
-10	1.21772	0.332761	0.0807504	0.0204412	0.00918484

Table 5–3: Cylindrical Bending Case – Spatially Sampled Intensity Rate Data

As can be seen, at the finer grid spacings, the proportional deterioration caused by the spatial discretisation was greater in the bending case, due to the previous greater accuracy; the deterioration for the toroidal case was absorbed into the error already present. While it might seem odd that for some cases the approximate intensity rate spatial derivatives yield better results than the analytic ones, it should be recalled that the results with the analytic derivatives are already imperfect due to the intrinsic spatial sampling of the velocity in the solution method. Thus errors in the intensity rate derivatives may contribute

Increment	0.4	0.2	0.1	0.05	0.0333333
True Velocity	10.5283	10.9440	11.1511	11.2544	11.2885
$\log_{10}\lambda$	Velocity Errors (rms)				
10	33.4801	90.8254	26.395	*	*
8	34.2588	110.035	21.9852	*	*
6	16.8352	21.4448	24.525	*	*
4	2.42365	2.44608	2.43186	*	*
2	0.812718	0.261943	0.0929608	*	*
0	0.828699	0.274659	0.102278	0.0577022	0.0575961
-2	0.828863	0.274799	0.102395	0.0574852	0.0589994
-4	0.828865	0.274800	0.102399	0.0575691	0.0591179
-6	0.828865	0.274800	0.102399	0.0575638	0.0583149
-8	0.828865	0.274800	0.102399	0.0575638	0.0583149
-10	0.828865	0.274800	0.102399	0.0575638	0.0583149

Table 5–4: Toroidal Stretching Case – Spatially Sampled Intensity Rate Data

Increment	120	60	30	15	10
True Velocity	5.20239	8.3905	9.87864	10.5991	10.8358
$\log_{10}\lambda$	Relative Velocity Errors				
10	-0.00000	0.00000	0.00004	*	*
8	-0.00028	0.00047	0.00383	*	*
6	-0.02053	0.01086	0.06237	*	*
4	2.83883	4.05158	3.35732	3.65581	3.16741
2	2.54526	3.15228	3.06773	3.05754	2.83738
0	2.52329	3.12580	3.04593	3.06707	2.85772
-2	2.52308	3.12565	3.04731	3.07245	2.86451
-4	2.52307	3.12567	3.04729	3.07220	2.86776
-6	2.52307	3.12567	3.04729	3.07222	2.86781
-8	2.52307	3.12567	3.04729	3.07218	2.86752
-10	2.52307	3.12567	3.04729	3.07224	2.86671

Table 5–5: Cylindrical Bending Case – Relative Error for Spatial Sampling

Increment	0.4	0.2	0.1	0.05	0.0333333
True Velocity	10.5283	10.9440	11.1511	11.2544	11.2885
$\log_{10}\lambda$	Relative Velocity Errors				
10	0.00001	0.00001	-0.00008	*	*
8	0.00050	0.00157	-0.00007	*	*
6	-0.01281	-0.00471	-0.00151	*	*
4	-0.10500	-0.04584	-0.01711	*	*
2	2.35106	2.25753	0.59180	*	*
0	2.33660	2.47449	0.88655	0.13482	0.03419
-2	2.33623	2.47452	0.88813	0.13834	0.03296
-4	2.33622	2.47449	0.88797	0.13758	0.03377
-6	2.33622	2.47448	0.88797	0.13733	0.03286
-8	2.33622	2.47448	0.88797	0.13733	0.03286
-10	2.33622	2.47448	0.88797	0.13733	0.03286

Table 5–6: Toroidal Stretching Case – Spatially Sampled Intensity Rate Data

to better overall results. It should be noted that this does not occur at the values of λ producing the best results.

It was from consideration of these results that the cases used for the subsequent sections, in which the results of adding various levels of noise will be given, were chosen. As the grids with more than 16×16 increments were computationally expensive, taking into account the requirement for repetition in the cases where noise is added, this was the grid size chosen. The results with this and finer grid sizes made $\lambda = 10^{-4}$ seem a plausible choice for the regularisation parameter, giving reasonable results in both the bending and stretching cases. It should be noted that this choice of λ is one in which there is little trade-off between the errors of the two terms in the regularisation, so the results in the following sections show that the method succeeds without the most favourable choice of λ.

5.3.3 Noisy Intensity Rate Input

This section contains the results of adding noise to the intensity rate input which was used as in the previous section in which the intensity rate spatial derivatives were found by finite differencing. The intensity rate at each sampled point in the image was corrupted by the addition of zero mean Gaussian noise with various standard deviations as given in the tables. The level of noise added may be compared against the rms intensity rates over the two regions, (calculated over the sampled points), which are 24.9819 for the bending case and 11.226 for the stretching case.

As previously stated, the results in the table were found from 100 trials of the calculation of the velocity field over the surface region, and are the mean (over the 100 trials), denoted μ, and the estimated standard deviation (over the 100 trials), denoted σ, of the rms velocity error over the region in an individual trial. Thus, the mean may be compared directly with the comparable result in the previous section, and the standard deviation used to judge the significance of any difference. In order to facilitate this, the corresponding noise-free result from the previous section is given below (as $\mu(0)$), as are the percentage difference between the mean and $\mu(0)$ and the standard deviation as a percentage of $\mu(0)$.

Noise Level	10	1	0.1	0.01	0
μ	2.84907	0.293033	0.0847693	0.0807079	0.0807503
σ	0.797412	0.0822883	0.00779806	0.000737633	0
$\frac{\mu-\mu(0)}{\mu(0)}\%$	3428.247	262.888	4.977	-0.053	0
$\frac{\sigma}{\mu(0)}\%$	987.503	101.905	9.65700	.913474	0

Table 5–7: Cylindrical Bending Case – Noise Added to Intensity Rate

When the results at the 10, 1 and 0.1 noise levels are considered, it can be seen that it is at about the 1 level at which the noise on the intensity rate produces errors of the same order of magnitude as the error in the original discretised solution. This is consistent with requiring intensity rates to an

Noise Level	10	1	0.1	0.01	0
μ	2.69729	0.289711	0.10596	0.102368	0.102399
σ	1.55016	0.14769	0.00834045	0.000806962	0
$\frac{\mu-\mu(0)}{\mu(0)}\%$	2534.098	182.924	3.478	-0.030	0
$\frac{\sigma}{\mu(0)}\%$	1513.84	144.23	8.14505	.788057	0

Table 5–8: Toroidal Stretching Case – Noise Added to Intensity Rate

accuracy greater than 5% that of the rms intensity rates in the cylindrical case.

The normally expected trend of decreasing input noise yielding decreasing errors was broken by the fact that adding errors of standard deviation 10^{-2} produced a mean result better than the zero noise case. However, with a sample of only 100 trials, making the assumption that the rms error in any one trial was itself a Gaussian variable would lead to the standard deviation of the mean rms error being 0.1 that of the individual rms error; hence the difference between the mean rms error in the noisy case and the rms error in the noise-free case was less than the standard deviation of that mean, and not significant.

It might perhaps appear surprising that it is ever possible for the noisy case to produce better results than the noise-free case, but it should be remembered that the noise-free case was not error-free, containing as it did the effect of the spatial discretisation of both the velocity (necessary for this method of solution) and the intensity rate (necessary for the plausible addition of the intensity rate noise). Thus, a small perturbation of any coefficient of the simultaneous linear equations which represented the discretised form of the problem could lead to either a better or a worse result, as could the effect of superposing such perturbations. A further test was conducted in which rather than adding the 10^{-2} noise to the intensity rate, the same perturbations were subtracted, and the results were that the mean rms error was greater than the noise free rms error, by an amount slightly greater than it had been less when the noise was added. Clearly, over an infinite set of trials, the subtraction of zero mean

Gaussian noise with a certain standard deviation should produce equivalent results to the addition of such noise, but over a finite set of trials, the result of adding or subtracting a sequence of perturbations to the same elements of the system will produce different results. Indeed if the system were to behave linearly in respect of the perturbations, the result of subtracting them would be the opposite of the result of adding them.

5.3.4 Noisy Velocity Input

This section contains the result of adding noise to the velocity input, with the calculation otherwise the same as in the section on the spatially sampled intensity rate. Each of the normal and two tangential velocity components at each sample point on the bounding contour were corrupted by zero mean Gaussian noise of various standard deviations as shown in the tables. The level of noise can be compared against the true rms velocity as given previously, 9.87864 in the cylindrical case and 11.1511 in the toroidal.

As in the previous section, the mean (μ) and estimated standard deviation (σ) of the rms errors over 100 trials is given, with the noise-free results reproduced for comparison.

Noise Level	1	0.1	0.01	0.001	0
μ	1.51159	0.170337	0.0819506	0.0807454	0.0807503
σ	0.146479	0.0150478	0.00204588	0.000209703	0
$\frac{\mu-\mu(0)}{\mu(0)}\%$	1771.931	110.943	1.486	-0.006	0
$\frac{\sigma}{\mu(0)}\%$	181.397	18.635	2.53359	.259693	0

Table 5–9: Cylindrical Bending Case – Noise Added to Bounding Contour Velocity

The level at which the boundary velocity noise introduces errors of the same order of magnitude as are already present in the noise-free case is that of 0.1, the level of the velocity error in the noise-free case, and about 1% of the rms velocity level. As in the trials involving noise on the intensity rate

Noise Level	1	0.1	0.01	0.001	0
μ	1.63759	0.194504	0.104094	0.102456	0.102399
σ	0.166132	0.0209998	0.00280365	0.000281474	0
$\frac{\mu-\mu(0)}{\mu(0)}\%$	1499.225	89.947	1.655	0.056	0
$\frac{\sigma}{\mu(0)}\%$	162.24	20.5078	2.73797	.27488	0

Table 5–10: Toroidal Stretching Case – Noise Added to Bounding Contour Velocity

input, the general trend of decreasing input noise leading to decreasing output velocity errors was broken, but the difference between the 10^{-3} and noise-free cylindrical bending case was not significant and the mean of the rms output velocity error was greater than the noise-free case when the same noise sequence was subtracted instead of added.

5.3.5 Noisy Light Source Input

The previous sections, in which the addition of noise has been considered, have been concerned with the addition of noise to the dynamic inputs, the intensity rate and the velocity, since in motion interpretation they constitute the primary inputs. In this and the following section, the addition of noise to the static inputs, the light source information and the shape information, is considered.

Whereas the manner in which the intensity rate constraint and the velocity on the bounding contour affect the solution for the velocity within the region is fairly direct, the effect of the light source and shape information is more obscure. The notional intensity gradient normal to the surface I_n, for example, is a quantity which is not directly measured but depends upon the lighting and surface shape. Further, unlike the dynamic inputs for which one may have a model of how they are expected to be measured (*e.g.* the intensity rate may be found by finite differencing of the intensity over a time period) the model of the origin of the light source and shape information is less clear.

The hypothetical dependence of the latter on the former, due to its presumed derivation from shape from shading may not be simple. However, to avoid being side-tracked into matters of static shape, the manner in which the noise was added was kept simple and this and the subsequent section may be regarded as constituting little more than evidence that the exact input of light source and shape information is not required.

The results in this section are from experiments in which each component of the light source vector was corrupted by zero mean Gaussian noise with various standard deviations as given in the tables. As has been done previously, the mean (μ) and estimated standard deviation (σ) of the rms errors over 100 trials is given, with the noise-free results reproduced for comparison.

Noise Level	0.1	0.01	0.001	0.0001	0
μ	6.16817	0.293352	0.0847588	0.0806657	0.0807503
σ	32.9089	0.18293	0.0152906	0.0015269	0
$\frac{\mu-\mu(0)}{\mu(0)}\%$	7538.572	263.283	4.964	-0.105	0
$\frac{\sigma}{\mu(0)}\%$	40753.9	226.538	18.9357	1.89089	0

Table 5–11: Cylindrical Bending Case – Noise Added to Light Source Vector

Noise Level	0.1	0.01	0.001	0.0001	0
μ	8.86852	0.224055	0.102844	0.102347	0.102399
σ	43.0294	0.446469	0.0059634	0.000584616	0
$\frac{\mu-\mu(0)}{\mu(0)}\%$	8560.748	118.806	0.435	-0.051	0
$\frac{\sigma}{\mu(0)}\%$	42021.3	436.009	5.82369	.57092	0

Table 5–12: Toroidal Stretching Case – Noise Added to Light Source Vector

As can be seen, there is more an order of magnitude change in the error when the accuracy to which the components of the light source vector are known changes from 0.01 to 0.1, so an accuracy of at least 0.01 seems to be required for any useful form of results.

5.3.6 Noisy Shape Information

It has already been commented that the lack of a model for the manner in which the shape information is derived renders a proper modelling of the results of adding noise impossible, as it is unclear to what the noise should be added directly and what should be noisy as a result of this. This occurs because it is presumed that "shape" information is available as required, not, for example, position information from which shape information may be estimated. In the case of shape from shading, the positional information depends in a complex fashion, potentially *via* shape parameters, on the static intensity image, which is perhaps the true source to which to add noise.

Consequently, the experiments reported in this section followed the simple principle of adding zero mean Gaussian noise (proportionally to the magnitude of the individual item affected) to all types of shape information (*e.g.* E and $J_{\alpha\alpha}$) at each sampled point. (At each sampled point, those quantities such as E which are by definition positive, were left unaltered if the noise which would have been added was such that they would have become negative). Thus, the relationships between various shape quantities will be destroyed, and the shape inputs will be mutually inconsistent. As stated previously, the results below demonstrate little more than that it is not necessary to measure all shape quantities perfectly, nor do they have to be mutually consistent.

Noise Level	0.1	0.01	0.001	0.0001	0
μ	1.34123	0.156437	0.0819607	0.0807724	0.0807503
σ	0.166911	0.018953	0.00233082	0.000233055	0
$\frac{\mu-\mu(0)}{\mu(0)}\%$	1560.960	93.729	1.499	0.027	0
$\frac{\sigma}{\mu(0)}\%$	206.7	23.4711	2.88645	.288612	0

Table 5–13: Cylindrical Bending Case – Noise Added to Shape Input

As can be seen there was approximately an order of magnitude difference between the results in which quantities were corrupted by 10% proportional noise and by 1% proportional noise, suggesting that at least some quantities

Noise Level	0.1	0.01	0.001	0.0001	0
μ	1.6255	0.172115	0.103467	0.102464	0.102399
σ	1.2544	0.0245984	0.00332718	0.000330939	0
$\frac{\mu-\mu(0)}{\mu(0)}\%$	1487.418	68.083	1.043	0.063	0
$\frac{\sigma}{\mu(0)}\%$	1225.01	24.0221	3.24923	.323186	0

Table 5–14: Toroidal Stretching Case – Noise Added to Shape Input

should be measured to the latter accuracy. Further investigation of this matter seems worthwhile only if the manner in which the quantities are derived is properly modelled, as discussed above.

5.3.7 Finite Time Increment

In this section results are given for cases in which, in contrast to the previous sections where the intensity rate and velocity at any point were calculated as differentials of the intensity and position respectively, the intensity rate and velocity were determined from the difference of the intensity and position over a finite time increment.

The results for a range of time increments are presented below. It should be noted that determining the velocities as differences over the finite time period yielded different "true" velocities to those in the differential case which is reproduced here as the zero time increment case for convenience of comparison. Analogously to the previous sections, the percentage difference between the rms velocity error (denoted $v(\Delta t)$ and that in the zero time increment case (denoted $v(0)$) will also be given.

Given that the intensity rate constraint, upon which the solution for the velocities was based, applies to the differential (or zero time increment) case, the results in the cylindrical bending case would be surprising if the effect of temporal discretisation were the only one factor influencing the results. However, the solution also includes the effect of spatial discretisation, and as time advances in the cylindrical bending case and the radius of the cylinder increases, the

Time Increment	1	0.1	0.01	0.001	0
True Velocity rms	7.81318	9.62472	9.85265	9.87603	9.87864
Velocity Error rms	1.74540	0.238154	0.0504663	0.0775867	0.0807503
$\frac{v(\Delta t)-v(0)}{v(0)}\%$	2061.478	194.926	-37.503	-3.918	0

Table 5–15: Cylindrical Bending Case – Finite Time Increments

Time Increment	1	0.1	0.01	0.001	0
True Velocity rms	10.9862	11.1343	11.1494	11.1509	11.1511
Velocity Error rms	1.42049	0.193287	0.108449	0.102954	0.102399
$\frac{v(\Delta t)-v(0)}{v(0)}\%$	1287.211	88.759	5.908	0.542	0

Table 5–16: Toroidal Stretching Case – Finite Time Increments

finite difference approximations to various spatial derivatives become more accurate, hence when this effect is not outweighed by that of the error introduced by the temporal discretisation, the results are better than in the differential case. Conversely, if the sign of the cylindrical motion is reversed, the effect of spatial discretisation becomes worse as time advances, and the results are worse than the differential case for all the finite time increments used above.

The results for the 0.1 and 0.01 time increment cases are base cases against which the results of adding noise to the intensity rate and to the velocity, as in the next two sections, can be judged, with less emphasis being placed on the cylindrical bending case. These were chosen as the time increments at which the effects of temporal discretisation were similar to and less than those of spatial discretisation.

5.3.8 Finite Time Increment — Noisy Intensity Rate Input

This section contains the result of adding noise to the intensity rate input, in the finite time increment case. Similarly to the previous case of adding noise to the intensity rate input, the intensity rate at each sampled point in the

image was corrupted by the addition of zero mean Gaussian noise with various standard deviations as given in the table. The level of noise added may be compared against the rms intensity rates over the two regions for the two time increments: 23.8526 and 24.864 for the cylindrical bending case, and 11.2046 and 11.2236 for the toroidal stretching case, for time increments 0.1 and 0.01 respectively.

The means ($\mu(0.1)$ and $\mu(0.01)$) and estimated standard deviations ($\sigma(0.1)$ and $\sigma(0.01)$) of the rms errors for the 0.1 and 0.01 time increment cases over 100 trials are given below, with the noise-free results reproduced for comparison. As in the previous sections where noise was added, the percentage difference of the mean rms error from the noise-free, *zero time increment* case ($\mu(0,0)$) and the standard deviation as a percentage of ($\mu(0,0)$) are also given.

Noise Level	10	1	0.1	0.01	0
$\mu(0.1)$	2.87003	0.377597	0.240708	0.238256	0.238154
$\sigma(0.1)$	0.788441	0.0762296	0.00771443	0.000780351	0
$\frac{\mu(0.1)-\mu(0,0)}{\mu(0,0)}\%$	3454.203	367.611	198.089	195.053	194.926
$\frac{\sigma}{\mu(0,0)}\%$	976.394	94.4016	9.55344	.966375	0
$\mu(0.01)$	2.84969	0.287449	0.0571383	0.050457	0.0504663
$\sigma(0.01)$	0.796416	0.0814307	0.00821704	0.000770641	0
$\frac{\mu(0.1)-\mu(0,0)}{\mu(0,0)}\%$	3429.015	255.973	-29.241	-37.515	-37.503
$\frac{\sigma}{\mu(0,0)}\%$	986.27	100.843	10.1759	.954351	0

Table 5–17: Cylindrical Bending Case – Finite Time Increment : Noise Added to Intensity Rate

When the results are considered in terms of the accuracy to which the input intensity must be determined to yield various accuracies of output, it appears that both the cases of the 0.1 time increment and the 1 error in intensity rate (corresponding to an error of order 0.1 in the intensity measurement) are points at which the output errors become worse by an order of magnitude when the intensity measurement does so. As these cases have errors of about 3%,

Noise Level	10	1	0.1	0.01	0
$\mu(0.1)$	2.70373	0.333877	0.194577	0.193235	0.193287
$\sigma(0.1)$	1.54365	0.141135	0.0138284	0.00144033	0
$\frac{\mu(0.1)-\mu(0,0)}{\mu(0,0)}\%$	2540.387	226.055	90.018	88.708	88.759
$\frac{\sigma}{\mu(0,0)}\%$	1507.49	137.829	13.5044	1.40659	0
$\mu(0.01)$	2.69767	0.291996	0.111694	0.108411	0.108449
$\sigma(0.01)$	1.54953	0.147095	0.00920861	0.000937391	0
$\frac{\mu(0.1)-\mu(0,0)}{\mu(0,0)}\%$	2534.469	185.155	9.077	5.871	5.908
$\frac{\sigma}{\mu(0,0)}\%$	1513.23	143.649	8.99287	.91543	0

Table 5–18: Toroidal Stretching Case – Finite Time Increment : Noise Added to Intensity Rate

and the underlying error due to the original discretisation is about 1%, it seems plausible to suggest that for both the cases considered, the method requires the intensity to be measured to an accuracy of order 0.1 for the resulting intensity rate to constitute an acceptable input to the method.

As the maximum intensity in both cases is almost 255, it is clear that this level of accuracy is much greater than that which can be achieved with standard camera technology, hence the physical reproduction of these experiments would be rather fruitless. This is, of course, not to deny that a case could be contrived in which the method could be demonstrated on real data, merely to suggest that such a case would indeed appear contrived in the colloquial sense.

5.4 Conclusion

It has been shown that for a general surface the application of a (square of) divergence regulariser to the intensity rate constraint (as found in Chapter 3) enables the estimation of surface velocities from the changing intensity image produced by a smoothly deforming Lambertian surface region of uniform albedo. Given perfect input data, the velocity estimates in the ideal continu-

ous case will match the true velocities for cases such as bending and uniform stretching motions. However, in practice the results are subject to discretisation errors in the solution and noise in the input.

In the case of the primary input, the intensity rate, the addition of noise in the synthesised test cases which excluded the effect of temporal discretisation showed that the instantaneous intensity rate would have to be measured to an accuracy better than 5% (of the rms intensity rate) to ensure that the errors induced by the noise were of a similar order of magnitude to those inherent in the solution. In the cases including the effect of temporal discretisation, the results suggested that intensity measurements 10 times more accurate than commonly available would be required to meet the same criterion. This demonstrates an important point regarding the design of "cameras" for computer vision – it would be desirable for many techniques to have temporal and / or spatial derivatives measured to similar levels of accuracy to that to which the intensity itself is measured. If one considers an analogy with the measurement of position in aircraft navigation systems, it is apparent that whereas in the case of the vision system for which intensity is the major input and its temporal derivatives are estimated by differencing, in the case of the aircraft navigation systems the temporal derivatives of position (velocity and acceleration) will also be measured. A more direct measurement of the derivatives has been employed in a VLSI implementation of motion estimation in (Mead, 1989) based upon the method of (Horn & Schunck, 1981).

As might be expected, errors in the boundary velocity produce errors of a similar order of magnitude in the output. The tests on the light source data suggest again that measurements of about 1% accuracy are required, and those on the shape data demonstrate some noise tolerance.

Chapter 6

Conclusions and Further Work

This thesis has considered the changing intensity pattern induced by a smoothly deforming constant albedo Lambertian surface illuminated by a distant point light source and viewed under orthographic projection. The two major original contributions are:

- The determination of the Intensity Rate Constraint which relates the temporal derivative of intensity at a fixed image point to the normal component of the three-dimensional velocity of the surface point corresponding to the image point.

- The development of a method using the Intensity Rate Constraint to recover the field of three-dimensional velocities over a region of a deforming surface by the application of regularisers chosen for the physical significance of the velocity fields which may be correctly recovered.

The Intensity Rate Constraint was derived in Chapter 3 for two cases. In the first the reflectivity, illumination and projection were as used in the rest of the thesis. The second was a similar more general configuration in which the reflectivity was a function (only) of the magnitude of the illumination and the angle between its direction and that of the surface normal, while the illumination could vary in both direction and magnitude with position, and the projection could be more general.

The method for estimating the velocity was presented in Chapter 4 for the reduced dimensional case, in which an analogy may be drawn with curve motion in the plane, and in Chapter 5 for the full dimensional case. Both cases, which used as inputs the intensity rate in the image, the static shape of the surface, the illumination and surface albedo, and the three-dimensional boundary velocities, were shown to work on synthetic data. The results of applying the method to synthetic data representing the full dimensional case showed that physical reproduction of the experiments conducted would require higher accuracy intensity measurements than are currently commonly available: 0.1 of a part in 255.

As the phenomenon considered includes the effect of the change in intensity induced at a surface point by the change in the angle between the surface normal and the illumination, it is a higher order effect than the usual phenomenon considered in motion estimation, in that, only the effect of the change in intensity at an image point due to the change in corresponding world point (which is itself presumed to be of unchanging intensity) is considered. Thus, as illustrated by the required accuracy mentioned above, this does not appear to be a practical approach to pursue further at present.

If intensity measurements of significantly greater accuracy became readily available, then it might be of interest to investigate the results obtainable in practice, which should be expected to show the effects of other major issues such as the matter of the practical realisability of the required reflectivity and illumination. In this event it would be worth considering the re-implementation of the numerical solution in the method, using the Finite Element approach to solve the partial differential equations, and thus have a more accurate treatment of the more arbitrarily shaped boundaries which might be encountered.

It might also be possible to pursue a more qualitative approach in which, for example, just the information that a change in intensity is related to a change in direction relative to the light source is used (an increase indicating that the angle between the surface normal and the direction of the light source has decreased and *vice versa*).

However, the direction of research which seems the most profitable to pursue further is one which would be required for any practical application of the method developed to interpret the Intensity Rate Constraint – the estimation of the three-dimensional velocity along a deforming space curve. While this is needed to provide the boundary velocity to the method developed in this thesis, it is also an interesting problem in its own right and this work concludes with a brief discussion of it.

6.1 Estimating the Velocity on a Deforming Space Curve

Similarly to the problem of estimating planar curve motion in the image which was discussed in Chapter 2 and the problem considered in this thesis, that of estimating the velocity of a deforming curve in 3D is subject to a form of aperture problem in that (except at endpoints) the tangential velocity is unobservable. This is immediately apparent upon consideration of a purely tangential motion in which the curve deforms onto itself.

A method for determining the depth, and translational and rotational velocities of a rigidly moving space curve from the (observable) normal component of image velocity, assuming monocular viewing under perspective projection, has recently been proposed in (Faugeras, 1990).

Whilst it is of interest to extend the type of motion to include deformation, there is no requirement to remain with monocular viewing. Indeed the application of stereo, potentially with more than two cameras (see *e.g.* (Ayache & Lustman, 1987)) to avoid problems with points where the curve is nearly tangential to an epipolar line, poses the problem in a form to which the method of regularisation may readily be applied.

When a point on a moving space curve is observed under a projection such that the projection ray is $\mathbf{p}$, the three-dimensional velocity component whose

projection is observable as the normal velocity component in the image will be the component in the following direction of the velocity at that point

$$\mathbf{O} = \mathbf{p} \times \mathbf{t} \tag{6.1}$$

where $\mathbf{t}$ is the (unit) tangent vector. Note that the issue of the sign of the direction will be left arbitrary as only the outline of the method is being sketched here.

This direction is in the *normal* plane, which is spanned by the normal and binormal vectors, $\mathbf{n}$ and $\mathbf{b}$. If we presume that the curve is viewed stereoscopically then its position $\mathbf{r}$ and all these vectors may be assumed known. Given this and the projections, each of the component three-dimensional velocities $\mathbf{v}_O$ in the directions $\mathbf{O}$ may be recovered from their corresponding observed normal components of the image velocities. As the method of stereo position determination relies upon the fact that the projection rays $\mathbf{p}$ corresponding to different cameras are not collinear, the corresponding directions $\mathbf{O}$ are also not collinear. Thus the normal and binormal components of the three-dimensional velocity are both recoverable, but the tangential component of velocity is unobservable.

The natural extension of the ideas presented in this thesis is to suggest that the velocity $\mathbf{v}$ be determined along an open curve, the full end-point velocities of which are known, by minimising the following integral along the curve with respect to $\mathbf{v}$

$$\int \left((v_n - v_{n_m})^2 + (v_b - v_{b_m})^2 + \lambda \left(\frac{\partial \mathbf{v}}{\partial s} \cdot \mathbf{t} \right)^2 \right) ds \tag{6.2}$$

where v_{n_m} and v_{b_m} are the measured normal and binormal components of velocity. Variations on this technique could include varying the weightings on the velocities in accordance with the presumed measurement accuracies or applying different regularisers, but the regulariser suggested is compatible with that used in this thesis, in that inextensible motions may be recovered exactly in the absence of noise, and uniform stretches as λ is decreased. As before, the regulariser has been chosen on the basis of the type of motion expected in the world rather than that expected in the image – the difference between the two being one of the points raised in (Faugeras, 1990).

6.2 A Final Comment

In retrospect the thesis appears to have taken something of a dimensional odyssey having started with the background in planar curve motion, dwelled on surface motion and ended with space curve motion. The underlying theme throughout these dimensional changes has been the derivation of theoretical constraints relating image changes to object motion and the application of regularisation terms intended to reflect assumptions about the type of object motion likely to be encountered, whether the object is an image or a world entity. It is perhaps this theme rather than any single result in the thesis which is most significant.

Appendix A

Numerical Implementation

A.1 Introduction

This appendix contains a description of the numerical implementation which was used to derive the results given in the thesis from the corresponding partial differential equations (PDEs). The explanatory emphasis will be placed on the general case addressed in chapter 5, rather than the cylindrical case of chapter 4, which is a subset of the general.

As the Intensity Rate Constraint which underlies the PDEs to be solved was derived in this work, rather than already being well-known, the numerical implementation was done in the most straightforward manner possible to facilitate the tracing of potential errors in the analysis through the implementation. The task of numerically solving coupled PDEs over a region given boundary conditions may be considered in two (related) parts: the discretisation of the PDEs and the solution of the resulting simultaneous equations. Thus this appendix is in four parts, the first two describing these aspects of the implementation, the third describing the general form of the implementation, including the manner in which the inputs to the method (the coefficients of the PDEs) were generated, and the output errors calculated, and the final part discussing the issue of verification.

Although this numerical implementation is based upon the requirement for solving the PDEs which arise as the result of applying the calculus of variations to the problem of formally minimising the integral arising from the method of

regularisation, it should be noted that an alternative approach starting from the minimisation of the integral could be considered. It would, for example, be possible, though presumably practically inefficient, to form a discrete approximation to the integral (as a summation of a discrete approximation to the integrand) and apply a general numerical minimisation method such as the well known simplex method of (Nelder & Mead, 1965) to the problem. Alternatively one might consider the approach of analytically minimising an approximation to the integral.

A.2 Discretisation of the Partial Differential Equations

The two main approaches to the numerical solution of linear PDEs, both of which result in a set of simultaneous linear equations to be solved, are *Finite Difference* (FD) methods (see *e.g.* (Smith, 1985)) and *Finite Element* (FE) methods (see *e.g.* (Davies, 1980)). Conceptually the two methods are very different in origin. The FD method is based upon approximating the PDEs at each point of a regular (usually rectangular) grid covering the region of interest by the use of Taylor's series expansions to approximate the differential terms, whereas the FE method is based upon approximating the solution to the PDEs over a set of (commonly triangular) sub-regions or elements (which sum together to form the region) by the use of trial functions with unknown parameters. However, in practice there may be less difference than might be expected from the difference in theoretical origin, as (Davies, 1980) comments (p. 263) of an example involving the application of both methods to Poisson's equation:

> The equivalence at interior points of the over-all equations has been known for some time, and it has often been said that one method is a special case of the other.

In general it appears that the FE methods are particularly superior in their treatment of the boundary, both in respect of the fact that the use of triangular elements permits a better geometric approximation of an arbitrary boundary shape than the rectangular grid associated with the finite difference approach, and in terms of the incorporation of the boundary conditions into the solution. However, in the case of this implementation, as the problem is synthesised, the boundary shape may be chosen to match the rectangular grid used in an FD approach, and the boundary conditions are of the simple (Dirichlet) form - the values of the functions sought are known on the boundary. Consequently the conceptually simpler FD method was used and its application will now be described further.

A.2.1 Application of the Finite Difference Method

As previously stated the FD method is based upon the approximation over a rectangular grid of the differentials of the function(s) sought by (truncated) Taylor's series. If the value of the smooth function $f(\alpha, \beta)$ is considered in the neighbourhood of the point (α_i, β_j), Taylor's theorem may be written (from (Stephenson, 1973)):

$$f(\alpha_i+h, \beta_j+k) = f(\alpha_i, \beta_j)+Df(\alpha_i, \beta_j)+\frac{1}{2!}D^2 f(\alpha_i, \beta_j)+...+\frac{1}{n!}D^n f(\alpha_i, \beta_j)+E_n$$

where

$$D^r f(\alpha_i, \beta_j) = \left(h\frac{\partial}{\partial \alpha} + k\frac{\partial}{\partial \beta}\right)^r f(\alpha_i, \beta_j)$$

and the error term

$$E_n = \frac{1}{(n+1)!}D^{n+1} f(\alpha_i + \theta h, \beta_j + \theta k)$$

for

$$0 < \theta < 1.$$

Letting this point (α_i, β_j) be part of a rectangular grid of width h in the α direction and k in the β direction, such that

$$\begin{aligned} \alpha_{i+1} &= \alpha_i + h \\ \beta_{j+1} &= \beta_j + k \end{aligned}$$

the Taylor series expansions up to second order for the value of $f(\alpha,\beta)$ at the eight neighbouring points of (α_i,β_j) (hereafter referred to simply as the point (i,j)) are:

$$\begin{aligned}
f(i-1,j-1) &= f(i,j) - h\frac{\partial f}{\partial \alpha} - k\frac{\partial f}{\partial \beta} + h^2\frac{\partial^2 f}{\partial \alpha^2} + hk\frac{\partial^2 f}{\partial \alpha \partial \beta} + k^2\frac{\partial^2 f}{\partial \beta^2} \\
f(i-1,j) &= f(i,j) - h\frac{\partial f}{\partial \alpha} + h^2\frac{\partial^2 f}{\partial \alpha^2} \\
f(i-1,j+1) &= f(i,j) - h\frac{\partial f}{\partial \alpha} + k\frac{\partial f}{\partial \beta} + h^2\frac{\partial^2 f}{\partial \alpha^2} - hk\frac{\partial^2 f}{\partial \alpha \partial \beta} + k^2\frac{\partial^2 f}{\partial \beta^2} \\
f(i,j-1) &= f(i,j) - k\frac{\partial f}{\partial \beta} + k^2\frac{\partial^2 f}{\partial \beta^2} \\
f(i,j+1) &= f(i,j) + k\frac{\partial f}{\partial \beta} + k^2\frac{\partial^2 f}{\partial \beta^2} \\
f(i+1,j-1) &= f(i,j) + h\frac{\partial f}{\partial \alpha} - k\frac{\partial f}{\partial \beta} + h^2\frac{\partial^2 f}{\partial \alpha^2} - hk\frac{\partial^2 f}{\partial \alpha \partial \beta} + k^2\frac{\partial^2 f}{\partial \beta^2} \\
f(i+1,j) &= f(i,j) + h\frac{\partial f}{\partial \alpha} + h^2\frac{\partial^2 f}{\partial \alpha^2} \\
f(i+1,j+1) &= f(i,j) + h\frac{\partial f}{\partial \alpha} + k\frac{\partial f}{\partial \beta} + h^2\frac{\partial^2 f}{\partial \alpha^2} + hk\frac{\partial^2 f}{\partial \alpha \partial \beta} + k^2\frac{\partial^2 f}{\partial \beta^2}
\end{aligned}$$

These (or higher order expansions) may be used to derive approximations to the first and second order partial differentials of $f(\alpha,\beta)$ at (i,j). For example, when $k = h$ the simplest approximations to the first and second order partial differentials are:

$$\begin{aligned}
\frac{\partial f}{\partial \alpha} &\approx \frac{1}{2h}\left(f(i+1,j) - f(i-1,j)\right) \\
\frac{\partial f}{\partial \beta} &\approx \frac{1}{2h}\left(f(i,j+1) - f(i,j-1)\right) \\
\frac{\partial^2 f}{\partial \alpha^2} &\approx \frac{1}{h^2}\left(f(i+1,j) - 2f(i,j) + f(i-1,j)\right) \\
\frac{\partial^2 f}{\partial \alpha \partial \beta} &\approx \frac{1}{4h^2}\left(f(i+1,j+1) - f(i+1,j-1)\right. \\
&\quad \left. -f(i-1,j+1) + f(i-1,j-1)\right) \\
\frac{\partial^2 f}{\partial \beta^2} &\approx \frac{1}{h^2}\left(f(i,j+1) - 2f(i,j) + f(i,j+1)\right)
\end{aligned}$$

In accordance with the objective of keeping the implementation as straightforward as possible, these expressions were used in this work to approximate the velocity PDEs. However, for the spatial derivatives of the intensity (when not calculated analytically) the following discretisations, as in (Brady et al,

1985), which are better in the presence of significant noise, but worse in the ideal case, were used:

$$\begin{aligned}\frac{\partial f}{\partial \alpha} &\approx \frac{1}{6h}\left(f(i+1,j-1)+f(i+1,j)+f(i+1,j+1)\right.\\ &\quad \left.-f(i-1,j-1)-f(i-1,j)-f(i-1,j+1)\right)\\ \frac{\partial f}{\partial \beta} &\approx \frac{1}{6h}\left(f(i-1,j+1)+f(i,j+1)+f(i+1,j+1)\right.\\ &\quad \left.-f(i-1,j-1)-f(i,j-1)-f(i+1,j-1)\right)\end{aligned}$$

The PDEs were approximated by forming, at each point at which the three components of the velocity were to be determined, three linear equations, one for each of the three PDEs, with the derivative terms approximated as above. These equations were put into matrix form by numbering the points in the $(m \times m)$ square grid (over which the solution is to be found) from 1 to m^2 such that, for the pth point, using integer arithmetic:

$$\begin{aligned} i &= p/m \\ j &= p \bmod m \end{aligned}$$

For a square grid, width h, as was used in the implementation, this yields the following three finite difference equations at point p:

$$\begin{aligned}
&\frac{1}{4h^2}v_{\beta(p-m-1)} + \left(\frac{1}{h^2} - \frac{1}{2h}\frac{A_\alpha}{A}\right)v_{\alpha(p-m)} + \left(-\frac{1}{2h}\frac{A_\beta}{A}\right)v_{\beta(p-m)} \\
&\qquad + \left(\frac{1}{2h}J\right)v_{n(p-m)} - \frac{1}{4h^2}v_{\beta(p-m+1)} \\
&+ \left(-\frac{2}{h^2} + \frac{A_{\alpha\alpha}A - A_\alpha^2}{A^2}\right)v_{\alpha p} + \left(\frac{A_{\alpha\beta}A - A_\alpha A_\beta}{A^2}\right)v_{\beta_p} + (-J_\alpha)\,v_{np} \\
&-\frac{1}{4h^2}v_{\beta(p+m-1)} + \left(\frac{1}{h^2} + \frac{1}{2h}\frac{A_\alpha}{A}\right)v_{\alpha(p+m)} + \left(\frac{1}{2h}\frac{A_\beta}{A}\right)v_{\beta(p+m)} \\
&\qquad + \left(-\frac{1}{2h}J\right)v_{n(p+m)} + \frac{1}{4h^2}v_{\beta(p+m+1)} = 0 \qquad \text{(A.1)} \\
&\qquad \frac{1}{4h^2}v_{\alpha(p-m-1)} - \frac{1}{4h^2}v_{\alpha(p-m+1)} \\
&+ \left(-\frac{1}{2h}\frac{A_\alpha}{A}\right)v_{\alpha(p-1)} + \left(\frac{1}{h^2} - \frac{1}{2h}\frac{A_\beta}{A}\right)v_{\beta(p-1)} + \left(\frac{1}{2h}J\right)v_{n(p-1)}
\end{aligned}$$

$$+\left(\frac{A_{\alpha\beta}A - A_\alpha A_\beta}{A^2}\right) v_{\alpha p} + \left(-\frac{2}{h^2} + \frac{A_{\beta\beta}A - A_\beta A_\beta}{A^2}\right) v_{\beta_p} + (-J_\beta)\, v_{np}$$

$$+\left(\frac{1}{2h}\frac{A_\alpha}{A}\right) v_{\alpha(p+1)} + \left(\frac{1}{h^2} + \frac{1}{2h}\frac{A_\beta}{A}\right) v_{\beta(p+1)} + \left(\frac{1}{2h}J\right) v_{n(p+1)}$$

$$-\frac{1}{4h^2} v_{\alpha(p+m-1)} + \frac{1}{4h^2} v_{\alpha(p+m+1)} = 0 \qquad \text{(A.2)}$$

$$-\frac{1}{4h^2} 2 l_\alpha l_\beta v_{n(p-m-1)} + \left(\frac{1}{2h}\lambda J\right) v_{\alpha(p-m)}$$

$$+\left(-\frac{1}{h^2} l_\alpha^2 - \frac{1}{2h}\left(-J l_n l_\alpha - l_\alpha \frac{\partial l_\alpha}{\partial \alpha} - l_\beta \frac{\partial l_\alpha}{\partial \beta}\right)\right) v_{n(p-m)} + \frac{1}{4h^2} 2 l_\alpha l_\beta v_{n(p-m+1)}$$

$$+\left(\frac{1}{2h}\lambda J\right) v_{\beta(p-1)} + \left(-\frac{1}{h^2} l_\beta^2 - \frac{1}{2h}\left(-J l_n l_\beta - l_\alpha \frac{\partial l_\beta}{\partial \alpha} - l_\beta \frac{\partial l_\beta}{\partial \beta}\right)\right) v_{n(p-1)}$$

$$+\left(-\lambda J \frac{A_\alpha}{A}\right) v_{\alpha p} + \left(-\lambda J \frac{A_\beta}{A}\right) v_{\beta_p}$$

$$+\left(\frac{2}{h^2} l_\alpha^2 + \frac{2}{h^2} l_\beta^2 - \left(\left(J l_n - I_n\right) I_n - \lambda J^2 + l_\alpha \frac{\partial I_n}{\partial \alpha} + l_\beta \frac{\partial I_n}{\partial \beta}\right)\right) v_{np}$$

$$+\left(-\frac{1}{2h}\lambda J\right) v_{\beta(p+1)} + \left(-\frac{1}{h^2} l_\beta^2 + \frac{1}{2h}\left(-J l_n l_\beta - l_\alpha \frac{\partial l_\beta}{\partial \alpha} - l_\beta \frac{\partial l_\beta}{\partial \beta}\right)\right) v_{n(p+1)}$$

$$+\frac{1}{4h^2} 2 l_\alpha l_\beta v_{n(p+m-1)} + \left(-\frac{1}{2h}\lambda J\right) v_{\alpha(p+m)}$$

$$+\left(-\frac{1}{h^2} l_\alpha^2 + \frac{1}{2h}\left(-J l_n l_\alpha - l_\alpha \frac{\partial l_\alpha}{\partial \alpha} - l b \frac{\partial l_\alpha}{\partial \beta}\right)\right) v_{n(p+m)}$$

$$-\frac{1}{4h^2} 2 l_\alpha l_\beta v_{n(p+m+1)} - \left(J l_n - I_n\right) \dot{I} - l_\alpha \frac{\partial \dot{I}}{\partial \alpha} - l_\beta \frac{\partial \dot{I}}{\partial \beta} = 0 \qquad \text{(A.3)}$$

These equations may then be put in the standard form

$$Ax = b$$

where x $(3m^2 \times 1)$ is the column vector of velocities to be found, with (assuming the vectors are numbered starting at 1 as is mathematically conventional – in the programs they were numbered from 0 in accordance with the convention in "C" (Kernighan & Ritchie, 1978) in which the programs were written):

$$v_{\alpha_p} = x_{3(p-1)+1}$$

$$v_{\beta_p} = x_{3(p-1)+2}$$

$$v_{n_p} = x_{3(p-1)+3}$$

A and b are arranged such that the three discretised PDEs at the pth point, from minimising the integral expression with respect to v_α, v_β and v_n, are represented by rows $3(p-1)+1$, $3(p-1)+2$ and $3(p-1)+3$ of A (and correspondingly $b_{3(p-1)+1}$, $b_{3(p-1)+2}$ and $b_{3(p-1)+3}$) respectively. It should be noted that although A is $3m^2 \times 3m^2$, most elements are zero, as the discretisation is such that in any row only the 27 elements, which correspond to the three velocity components at the point to which this row applies and its eight neighbours can be non-zero; thus the matrix is sparse and banded with maximum upper and lower bandwidths of $3m+5$. This leads to the matter of solving the simultaneous linear equations.

A.3 Solution of the Simultaneous Linear Equations

The task of finding a numerical solution to the three PDEs has now been reduced to that of solving a set of simultaneous linear equations in the form

$$Ax = b$$

where A is known to be sparse (and banded). The standard method for solving such problems in the absence of any knowledge about the form of the matrix A (other than the assumption that it is non-singular so that a unique solution exists) is *Gaussian Elimination with partial pivoting*, in which the matrix A is factorised into a lower triangular matrix L, an upper triangular matrix U and a pivotal matrix P which represents the row interchanges performed during the factorisation to improve stability of the solution. The resulting factorised form enables a straightforward solution for x. The process may be briefly represented as below, for the details of each stage see *e.g.* (Golub & Van Loan, 1986).

Solve for P,L,U:	$PLU = A$
Solve for y:	$PLy = b$
Solve for x:	$Ux = y$

The implementation in this work was based on a translation into C of a FORTRAN implementation from (Conte & deBoor, 1981) using "scaled" partial pivoting, in which, during the LU decomposition, the pivot row is chosen to be that with the pivot element of greatest magnitude relative to the largest element in the same row. As has been mentioned, this method takes no account of any special structure of the matrix; thus it is more general than is required for this problem and correspondingly computationally more expensive. The algorithm for the decomposition may be briefly expressed thus:

For $k \leftarrow 1$ to $n-1$
  Determine pivot row p from $(k+1, \ldots, n)$
  Swap rows k and p of A
  For $i \leftarrow k+1$ to n
    $a_{ik} \leftarrow a_{ik}/a_{kk}$
    For $j \leftarrow k+1$ to n
      $a_{ij} \leftarrow a_{ij} - a_{ik} * a_{kj}$

This decomposition of an $n \times n$ matrix may be implemented such that it requires (ignoring terms of lower order) $\frac{1}{3}n^3$ *flops*, where the flop, due to C.B.Moler and used in (Golub & Van Loan, 1986), is defined as the computational expense of the following floating point computation:

$$s \leftarrow s + a_{ij} * b_j$$

(As the computational expense of the subsequent solution for x is of order n^2, it will be ignored in the following discussion). It should perhaps be observed that this is not the lowest order of complexity for a general LU decomposition, *e.g.* (Aho et al, 1974) comments regarding a recursive method:

n simultaneous equations can be solved in $O(n^{2.81})$ steps.

Solving the simultaneous linear equations resulting from the discretisation over an $m \times m$ grid using the decomposition algorithm above requires $9m^6$ flops most of which will be wasted as the matrix is appreciably sparse. This expense can be significantly and straightforwardly reduced by taking account of the banded nature of the matrix. The application of a partial pivoting decomposition method to a matrix of upper bandwidth s is known to result in a matrix U with upper bandwidth no greater than $2s$ (Golub & Van Loan, 1986), and the computation may be reduced accordingly by searching only up to s rows for the pivot row (a minor effect) and by eliminating only up to s rows ahead for only $2s$ columns as below:

For $k \leftarrow 1$ to $n-1$
    Determine pivot row p from $(k+1, \ldots, min(k+s, n)$
    Swap rows k and p of A
    For $i \leftarrow k+1$ to $min(k+s, n)$
        $a_{ik} \leftarrow a_{ik}/a_{kk}$
        For $j \leftarrow k+1$ to $min(k+2s, n)$
            $a_{ij} \leftarrow a_{ij} - a_{ik} * a_{kj}$

This reduces the amount of computation required to $2ns^2$ for a general banded matrix, or (ignoring terms of lower order) $54m^4$ in the case of the solution of the discretised PDEs considered here. A less significant reduction in both the computation required when solving for x given the decomposition and the storage required can also be achieved because u_{ij} is zero when $j > i+2s$. Unfortunately, in the presence of pivoting there is no such simple result relating to the bandwidth of L, as may be seen by considering the possibility that the first row of A is repeatedly swapped in the pivoting process to become the nth row of L which could thus have all non-zero elements.

The storage required by the decomposition is also a significant burden – the full matrix contains $9m^4$ elements whilst the initial matrix has only $81m^2$

non-zero elements. In practice this meant that, using single precision floating point elements, the case of the grid with 16 increments (*i.e.* solving over a 15×15 grid knowing the boundary velocities on the 17×17 grid) was the largest which could be solved with a full two-dimensional array representing the whole of the matrix, without the memory requirement causing a drastic increase in the time taken. In the case of much larger grids the only solution methods which would be sufficiently compact in terms of storage would be those in which the sparseness of the matrix A is fully exploited, *i.e.* iterative schemes such as conjugate gradient methods or simultaneous over relaxation, (see *e.g.* (Press et al, 1988)).

However, the direct decomposition method may be trivially modified to potentially cope with slightly larger grids, such as those with 32 and 48 increments. In practice, the nature of the matrix is such that the row interchanges in the pivoting process are usually fairly local. If the interchanges are such that a row is never moved further than s from its initial position, the resulting decomposition will have maximum upper and lower bandwidths of $2s$ and thus only $n(4s+1)$ elements, or in the specific case considered, $36m^3 + 63m^2$ elements need be stored. Unfortunately the maximum interchange distance in the algorithm is not known *a priori*; thus it is not guaranteed that such a compact storage will permit the pivoting desired. However, for most cases of significant interest, this approach was satisfactory with a check employed on the pivoting to deem the equations as insoluble by the method if the pivoting criterion would result in a row being further than s from its original position. (See sections 5.3.1 and 5.3.2 for results obtained with this method, and asterisked instances where it did not succeed).

Another approach which was also tried was to use a pivotless decomposition, which requires only ns^2 ($27m^4$) flops (Golub & Van Loan, 1986) and $n(2s+1)$ or ($18m^3 + 33m^2$) elements of storage for the matrix. However, the results on the large grids (47×47) on which the pivoting method also worked showed a significant difference; so iterative improvement was applied to verify that the

problem lay in the pivotless approach. A step of the iteration used to improve the estimate $\hat{x}$ is shown below:

$$r \leftarrow b - A\hat{x} \text{ (done in double precision)}$$

Solve for y: $PLy = r$

Solve for e: $Ue = y$

$$\hat{x} \leftarrow \hat{x} + e$$

On the cases tried, two steps of this iterative improvement substantially reduced the residual r resulting in solutions close to those of the pivoting solution – subsequent steps made little change. However, for the toroidal stretching instance on the 48 increment grid, these solutions were not always closer to the true continuous solution. The first step of the pivotless approach produced results which were, for some cases, significantly more accurate as judged against the continuous solution (with errors reduced by a factor of 2); thus it appears that the limits of the approach have been reached.

A.4 General Implementation

The previous sections have described how the partial differential equations to be solved for the velocity field were discretised and how the resulting simultaneous linear equations were solved. This section describes the rest of the implementation, the general form of which was as below (in the toroidal case – different input parameters apply in the cylindrical case):

Read in the input data $1, \rho, a, r, v_x, \omega, \theta, \lambda$,

grid parameters, noise parameters, time increment dt, number of cases

Save 1

Print out the inputs

Set to zero all variables used to collect error information over multiple cases

For each test case do:

Restore 1

Set up the compact form of the matrix – if using this

Set to zero all variables used to collect information over a single case

Set the matrix A and vector b to zero

For every point on the $(m+2) \times (m+2)$ grid (including the boundary) do:

Determine $\alpha, \beta, \sin\alpha, \cos\alpha, \sin\beta, \cos\beta, \sin\omega t, \cos\omega t$

Determine x, y, z

Determine $\mathbf{r}_\alpha, \mathbf{r}_\beta, \mathbf{n}, E, E_\alpha, E_\beta, G_\alpha, G_\beta, A, A_\alpha, A_\beta, A_{\alpha\alpha}, A_{\alpha\beta}, A_{\beta\beta}$,

$J, J_\alpha, J_\beta, L, N$

Determine $l_\alpha, l_\beta, l_n, \frac{\partial l_\alpha}{\partial \alpha}, \frac{\partial l_\alpha}{\partial \beta}, \frac{\partial l_\beta}{\partial \alpha}, \frac{\partial l_\beta}{\partial \beta}$

If $dt = 0$ then

Determine continuous v_α, v_β, v_n

else

Determine x, y, z at time dt

Determine v_α, v_β, v_n from the differences in x, y, z / dt

If $dt = 0$ then

Determine continuous $\dot{I}$

else

Determine I

Determine α, β corresponding to image point x, z at time dt

Determine I at time dt – check that $\mathbf{l} \cdot \mathbf{n} > 0$

Determine $\dot{I}$ = change in I / dt

Determine $\frac{\partial \dot{I}}{\partial \alpha}, \frac{\partial \dot{I}}{\partial \beta}$ – if continuous time and spatial derivative case

Add noise to components of $\mathbf{l}$

For every point on the $(m+2) \times (m+2)$ grid (including the boundary) do:

Determine $\alpha, \beta, \sin\alpha, \cos\alpha, \sin\beta, \cos\beta, \sin\omega t, \cos\omega t$

Add proportional noise to $\mathbf{r}_\alpha, \mathbf{r}_\beta, \mathbf{n}, E, E_\alpha, E_\beta, G_\alpha, G_\beta$,

$A, A_\alpha, A_\beta, A_{\alpha\alpha}, A_{\alpha\beta}, A_{\beta\beta}, J, J_\alpha, J_\beta, L, N$

Determine $l_\alpha, l_\beta, l_n, \frac{\partial l_\alpha}{\partial \alpha}, \frac{\partial l_\alpha}{\partial \beta}, \frac{\partial l_\beta}{\partial \alpha}, \frac{\partial l_\beta}{\partial \beta}$

Determine $I_n, \frac{\partial I_n}{\partial \alpha}, \frac{\partial I_n}{\partial \beta}$

For every point on the $(m+2) \times (m+2)$ grid (including the boundary) do:

Add noise to $\dot{I}$

For every point on the $(m+2) \times (m+2)$ grid (including the boundary) do:

Add noise to velocities

For every point on the $m \times m$ grid (excluding the boundary) do:

Determine finite difference approximation to $\frac{\partial \dot{I}}{\partial \alpha}, \frac{\partial \dot{I}}{\partial \beta}$ – unless using continuous $\frac{\partial \dot{I}}{\partial \alpha}, \frac{\partial \dot{I}}{\partial \beta}$

For every point on the $m \times m$ grid (excluding the boundary) do:

Determine $\alpha, \beta, \sin\alpha, \cos\alpha, \sin\beta, \cos\beta, \sin\omega t, \cos\omega t$

Determine $l_\alpha, l_\beta, l_n, \frac{\partial l_\alpha}{\partial \alpha}, \frac{\partial l_\alpha}{\partial \beta}, \frac{\partial l_\beta}{\partial \alpha}, \frac{\partial l_\beta}{\partial \beta}$

Set up the 3 rows of the matrix A and vector b for this point

Solve the simultaneous linear equations $Ax = b$ for x

Set to zero variables used for information on rms velocity errors for this case

For every point on the $m \times m$ grid (excluding the boundary) do:

Determine estimated v_x, v_y, v_z from estimated v_α, v_β, v_n

Determine $\alpha, \beta, \sin\alpha, \cos\alpha, \sin\beta, \cos\beta, \sin\omega t, \cos\omega t$

Determine $\mathbf{r}_\alpha, \mathbf{r}_\beta, \mathbf{n}, E, E_\alpha, E_\beta, G_\alpha, G_\beta,$

$A, A_\alpha, A_\beta, A_{\alpha\alpha}, A_{\alpha\beta}, A_{\beta\beta}, J, J_\alpha, J_\beta, L, N$

Determine true v_x, v_y, v_z from true v_α, v_β, v_n

If first test case then

Output true and estimated v_x, v_y, v_z for possible display

Sum the square of the velocity error – the difference between the true and estimated v_x, v_y, v_z

Determine the root mean square error for this case

Sum the rms errors over the cases

Sum the squares of the rms errors over the cases

Determine the mean (μ) of the rms errors over the cases

Determine the standard deviation (σ) of the rms errors over the cases

Output μ, σ

Exit

Where possible, the values of variables were determined using the known analytic expression for them. However, in the discrete time case, the α and β

of the point on the expanded torus which corresponds to the image grid point was found by a two-dimensional Newton-Raphson search (see *e.g.* (Press et al, 1988)) started from the initial ($t = 0$) α and β. Other than this, and the effect of spatial discretisation in the discrete approximations to $\frac{\partial \dot{I}}{\partial \alpha}$ and $\frac{\partial \dot{I}}{\partial \beta}$, all the coefficients in the noise-free partial differential equations are derived from the analytic expressions.

A.5 Verification

The implementation performs sufficient computation to render hand-checking of a complete program run impractical, and relies upon many analytic expressions generated for the implementation in the absence of a source to check them against. Consequently, proper verification of the implementation to the standards applicable to industrial systems was not possible.

At the lower level it was possible to perform a limited check on program sections or functions such as the solution of the simultaneous equations by using inputs with known expected outputs. In the case of quantities for which the expressions were derived for the implementation some examples were checked by hand. However, the major part of the limited verification took the form of self-consistency checks, as will be described.

The expressions used to generate, at each point of the grid, the coefficients of the PDEs in the implementation, were also used in a program in which the velocity components and their spatial derivatives were calculated from analytic expressions. This enabled the near-cancellation of the terms in the PDEs which theoretically should sum to zero to be checked. In the case of the PDE derived from minimising the integral with respect to v_n for the toroidal expansion, the near-cancellation only occurs when λ is sufficiently small, as should be expected, as otherwise the "true" velocities are not the solution of the PDE.

The generation of the simultaneous equations representing the discretised continuous time PDEs was checked for consistency by observing the reduction

with the square of the grid increment of the rms discretisation error in each of the linear equations, *i.e.* measuring the rms value (normalising the error in the v_α equation by $\sqrt{E}$ *etc.*) of the vector:

$$Ax - b$$

The calculations of the discrete time velocity components and intensity rate were checked for convergence to the continuous time results.

Finally, there were the output estimated velocity components which, in the absence of noise, generally followed the expected trends of improvement with decreasing grid size, with the notable exception of the 48 increment grid for the toroidal expansion case. As has already been pointed out this case shows the limitations of the approach adopted.

Appendix B

Output Error Distributions

The results of applying the surface motion interpretation technique to noisy data were given in Chapter 5 in the form of the mean (μ) and standard deviation (σ) of the rms velocity output errors over 100 instances of each noisy test case. This was the only information given about the output error distribution. Thus the interpretation of the significance of the differences between the means of the noisy data and the noise-free case, in terms of an underlying assumption that the output errors were Gaussian, was not fully supported in that chapter.

Strictly it cannot be the case that these errors are Gaussian as the individual rms errors are lower bounded by zero. However, examination of the output error distributions at the lowest noise values used shows that they are near Gaussian, whereas those of the higher noise values are not. Given the relative magnitudes of μ and σ, the lower bound of zero does affect the distribution at the high noise level, but this is not the only effect present. The non-linearity (evident in some cases) of the variation of σ with input noise level and the change in shape of the error distribution suggests that the process from input noise to output error may become non-linear at high levels of noise. Given that the numerical solution method involves solving several simultaneous linear equations, this change from linear to non-linear is reasonable.

This general trend will be illustrated by giving histograms of the output errors for a few examples. These will show the number of cases (which as it is out of 100 may be regarded as a percentage!) which had output errors varying

from the mean by so many standard deviations. (The results more than 5σ from the mean were lumped into those between 4 and 5 σ. The examples given will be that of the cylindrical bending with noise (levels 0.01 and 10) added to the intensity rate, and the toroidal stretching with noise (levels 0.0001 and 0.1) added to the light source information. These represent both forms of motion and types of noise straightforward and indirect in their effect.

A more quantitative comparison with a Gaussian distribution may be conducted by considering the *skew* and *kurtosis* of the distributions, where the skew and kurtosis of the n measurements v_i is defined as (from (Press et al, 1988)):

$$skew = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{v_i - \mu}{\sigma}\right)^3$$

$$kurtosis = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{v_i - \mu}{\sigma}\right)^4 - 3$$

For a Gaussian distribution, the skew estimator has a standard deviation of $\sqrt{6/n}$ and the kurtosis estimator a standard deviation of $\sqrt{24/n}$. Thus for $n = 100$ a skew of magnitude less than 0.49 or a kurtosis of magnitude less than 0.98 is not significant at the 95% level.

The skew and kurtosis of the distributions which will be histogrammed are given in Table B.

	Cylindrical Bending With Intensity Noise		Toroidal Stretching With Light Source Noise	
Noise	0.01	10	0.0001	0.1
Skew	.43	.88	0.01	8.49
Kurtosis	.19	.37	.54	75.97

As can be seen from the table and the figures, while even the low noise level error distributions do not appear ideally Gaussian (some results were more so, *e.g.* the toroidal stretching case with 0.01 intensity rate noise had skew and kurtosis of 0.10 and 0.23 respectively), the high noise level error distributions are appreciably less so.

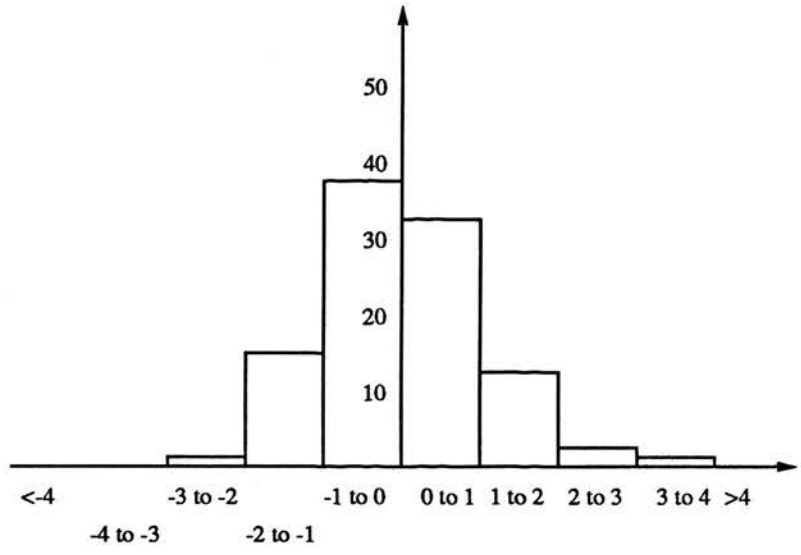

Figure B–1: Error Histogram: Cylinder Intensity Rate Noise (0.01)

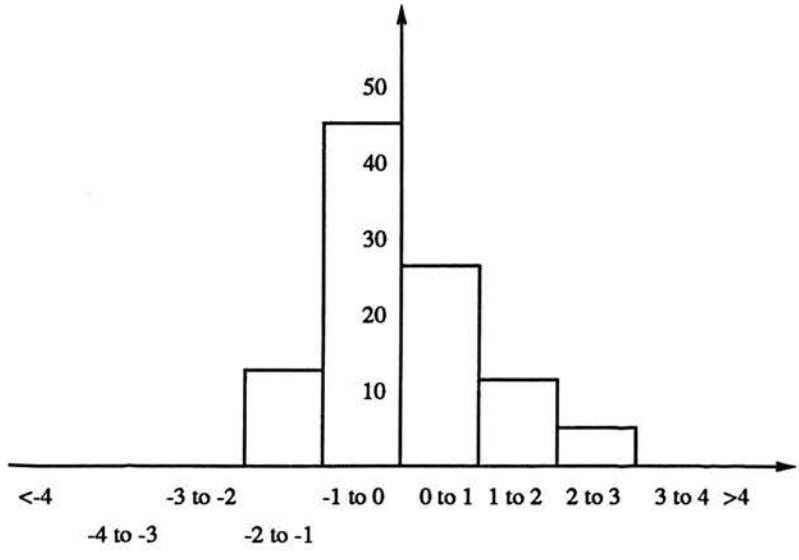

Figure B–2: Error Histogram: Cylinder Intensity Rate Noise (10)

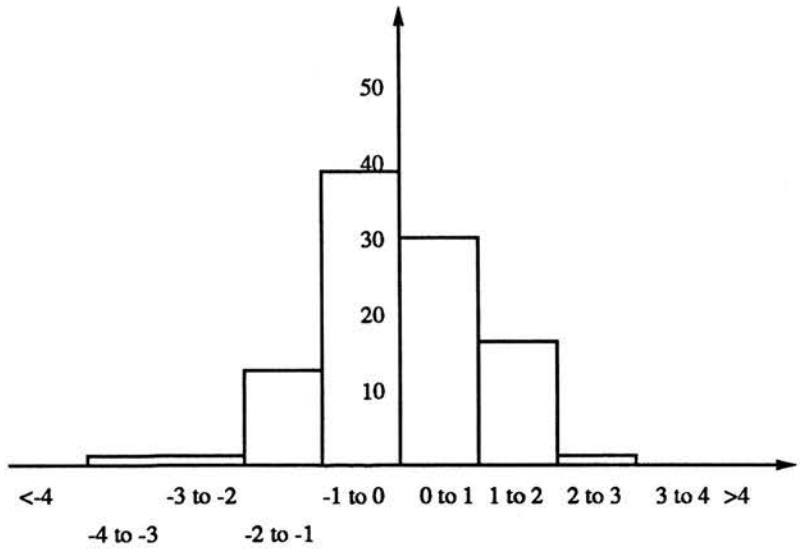

Figure B–3: Error Histogram: Torus Light Source Noise (0.0001)

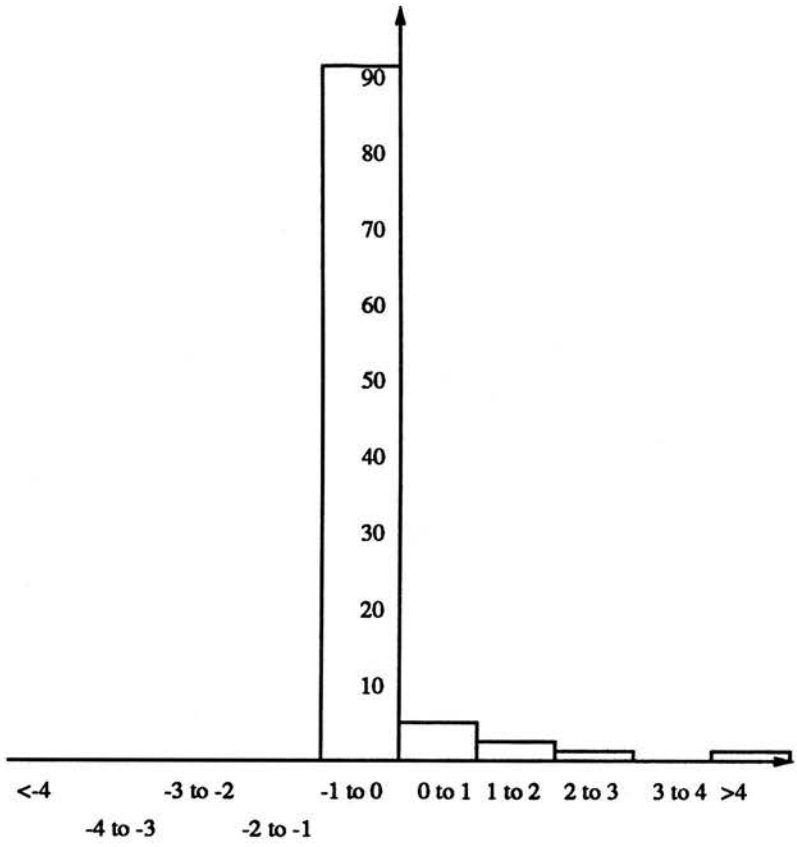

Figure B–4: Error Histogram: Torus Light Source Noise (0.1)

Glossary of Symbols

Bibliography

Aho, A.V., Hopcroft, J.E., and Ullman, J.D. (1974). *The Design and Analysis of Computer Algorithms.* Addison-Wesley.

Aloimonos, J. and Rigoutsos, I. (1986). Determining the 3-D motion of a rigid surface patch without correspondence under perspective projection: (i) planar surfaces, (ii) curved surfaces. In *Proceedings of the Fifth National Conference on Artificial Intelligence*, pages 681–688.

Ayache, N. and Lustman, F. (1987). Fast and reliable passive trinocular stereovision. In *Proceedings of the First International Conference on Computer Vision*, pages 422–427.

Bertero, M., Poggio, T., and Torre, V. (1987). *Ill-posed problems in early vision.* AI Memo 924, MIT.

Blake, A., Zisserman, A., and Knowles, G. (1985). Surface descriptions from stereo and shading. *Image and Vision Computing*, 3(4):183–191.

Brady, J.M. (1987). Seeds of perception. In *Proceedings of the Third Alvey Vision Conference*,pages 259–265.

Brady, M., Ponce, J., Yuille, A., and Asada, H. (1985). *Describing Surfaces.* AI Memo 822, MIT.

Brooks, M.J. and Horn, B.K.P. (1985). Shape and source from shading. In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, pages 932–936.

Brooks, R.A. (1981). Symbolic reasoning among 3-D models and 2-D images. *Artificial Intelligence*, 17:285–348.

Buxton, H. and Williams, N. (1986). Application of a fast parallel algorithm for the extraction and interpretation of optical flow. In *Proceedings of the Seventh European Conference on Artificial Intelligence*, pages 539–545.

Cameron-Jones, R.M. (1988). Visual interpretation of cylindrical deformation - a sideways look at contour motion! In *Proceedings of the Fourth Alvey Vision Conference*, pages 135–140.

Chen, S. and Penna, M. (1986). Shape and motion of nonrigid bodies. *Computer Vision Graphics and Image Processing*, 36:175–207.

Conte, S.D. and de Boor, C. (1981). *Elementary Numerical Analysis.* McGraw-Hill.

Cornelius, N. and Kanade, T. (1983). Adapting optical flow to measure object motion in reflectance and x-ray image sequences. In *SIGART Workshop on Motion Representation and Perception, Toronto*, pages 50–58.

Davies, A.J. (1980). *The Finite Element Method: A First Approach.* Oxford University Press.

D'Haeyer, J. (1986). Determining motion of image curves from local pattern changes. *Computer Vision, Graphics, and Image Processing*, 34:166–188.

do Carmo, M.P. (1976). *Differential Geometry of Curves and Surfaces.* Prentice-Hall.

Faugeras, O. (1990). On the motion of 3D curves and its relationship to optical flow. In *Proceedings of the First European Conference on Computer Vision*, pages 107–117.

Fisher, R.B. (1989). *From Surfaces To Objects: Computer Vision and Three Dimensional Scene Analysis.* John Wiley and Sons.

Gang, X. and Saburo, T. (1989). Three-dimensional interpretation of quadrilaterals. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, pages 1610–1616.

Gennert, M.A. and Yuille, A.L. (1988). Determining the optimal weights in multiple objective function optimization. In *Proceedings of the Second International Conference on Computer Vision*, pages 87–89.

Golub, G.H. and Van Loan, C.F. (1986). *Matrix Computations.* North Oxford Academic.

Gong, S. (1988). Improved local flow. In *Proceedings of the fourth Alvey Vision Conference*, pages 129–134.

Grzywacz, N.M. and Hildreth, E.C. (1987). Incremental rigidity scheme for recovering structure from motion : position-based versus velocity-based formulations. *Journal of the Optical Society of America A*, 4(3):503–518.

Hanson, A.R. and Riseman, E.M. (1978). Visions : a computer system for interpreting scenes. In Hanson, A.R. and Riseman, E.M., (eds.), *Computer Vision Systems*, pages 303–333, Academic Press.

Hildreth, E.C. (1984). *The Measurement of Visual Motion.* MIT Press.

Horn, B.K.P. and Schunck, B.G. (1981). Determining optical flow. *Artificial Intelligence*, 17:185–203.

Horn, B.K.P. (1986). *Robot Vision.* MIT Press.

Horn, B.K.P., Woodham, R.J., and Silver, W.M. (1978). *Determining Shape and Reflectance Using Multiple Images.* AI Memo 490, MIT.

Ikeuchi, K. and Horn, B.K.P. (1981). Numerical shape from shading and occluding boundaries. *Artificial Intelligence*, 141–184.

Jansson, G. and Johansson, G. (1973). Visual perception of bending motion. *Perception*, 2:321–326.

Kernighan, B.W. and Ritchie, D.M. (1978). *The C Programming Language.* Prentice-Hall.

Koenderink, J.J. and van Doorn, A.J. (1986). Depth and shape from differential perspective in the presence of bending deformations. *Journal of the Optical Society of America A*, 3(2):242–249.

Koenderink, J.J. (1990). *Solid Shape.* MIT Press.

Longuet-Higgins, H.C. and Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B*, 208:385–397.

Longuet-Higgins, H.C. (1981). A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–135.

Longuet-Higgins, H.C. (1984). The visual ambiguity of a moving plane. *Proceedings of the Royal Society of London B*, 223:165–175.

Marr, D. (1982). *Vision.* Freeman.

McIvor, A.M. (1990). A test of camera noise models. In *Proceedings of the British Machine Vision Conference*, pages 355–359.

Mead, C.A. (1989). *Analog VLSI and Neural Systems.* Addison-Wesley.

Nagel, H.-H. (1987). On the estimation of optical flow: relations between different approaches and some new results. *Artificial Intelligence*, 33:299–324.

Negahdaripour, S. and Horn, B.K.P. (1985). Determining 3-D motion of planar objects from image brightness patterns. In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, pages 898–901.

Nelder, J.A. and Mead, R. (1965). A simplex method for function minimisation. *Computer Journal*, 7:308–313.

Pentland, A. and Williams, J. (1989). Perception of non-rigid motion: inference of shape, material and force. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, pages 1565–1570.

Pentland, A.P. (1984). Local shading analysis. *IEEE Transactions on Pattern Matching and Machine Intelligence*, 170–187.

Pickett, M.S. (1984). *Graphics applications in robotics : modeling objects and tasks.* Submitted to SIGGRAPH 84 from General Motors Research Lab, Warren, Michigan.

Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T. (1988). *Numerical Recipes in C: The Art of Scientific Computing.* Cambridge University Press.

Shulman, D. and Aloimonos, J. (1988). (non-)rigid motion interpretation: a regularised approach. *Proceedings of the Royal Society of London B*, 233:217–234.

Smith, G.D. (1985). *Numerical Solution of Partial Differential Equations: Finite Difference Methods.* Oxford University Press.

Stephenson, G. (1973). *Mathematical Methods for Science Students.* Longman.

Subbarao, M. (1988). *Interpretation of Visual Motion: A Computational Study.* Pitman.

Terzopoulos, D., Witkin, A., and Kass, M. (1987). Symmetry-seeking models for 3D object reconstruction. In *First International Conference on Computer Vision*, pages 269–276.

Tikhonov, A.N. and Arsenin, V.Y. (1977). *Solutions of Ill-Posed Problems.* Winston and Sons.

Ullman, S. (1979). *The Interpretation of Visual Motion.* MIT press.

Ullman, S. (1984). Maximising rigidity : the incremental recovery of 3-D structure from rigid and nonrigid motion. *Perception*, 13:255–274.

Verri, A. and Poggio, T. (1987). Against quantitative optical flow. In *First International Conference on Computer Vision*, pages 171–180.

Verri, A., Girosi, F., and Torre, V. (1989). Mathematical properties of the two-dimensional motion field: from singular points to motion parameters. *Journal of the Optical Society of America A*, 6 No.5:698–712.

Wahba, G. (1980). *Ill-posed Problems: Numerical and Statistical Methods for Mildly, Moderately and Severely Ill-Posed Problems With Noisy Data.* Technical Report 595, University of Wisconsin, Madison.

Waxman, A.M. and Wohn, K. (1986). Image flow theory: a framework for 3-D inference from time-varying imagery. In Brown, C.M., (ed.), *Advances in Computer Vision*, pages 165–224, Lawrence Erlbaum.

Weatherburn, C.E. (1930). *Differential Geometry of Three Dimensions.* Volume 2, Cambridge University Press.

Weatherburn, C.E. (1931). *Differential Geometry of Three Dimensions.* Volume 1, Cambridge University Press.

Wolff, L.B. (1989). Shape understanding from Lambertian photometric flow fields. In *Proceedings of the IEEE CS Conference on Computer Vision and Pattern Recognition*, pages 46–52.

Woodham, R.J. (1978). *Photometric Stereo.* AI Memo 479, MIT.

Yuille, A.L. (1985). The smoothest velocity field and token matching schemes. In T.O'Shea, (ed.), *Advances in Artificial Intelligence*, pages 127–136, Elsevier Science Publishers.